

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 15:25:26 / Search time 40.3032 Seconds
(without alignments)
2847.400 Million cell updates/sec

Title: US-10-009-945-2
Perfect score: 3884
Sequence: 1 GGSSIKRLTVLCARLNANK.....EXLYEKULAVETGPAVE 723

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues 1107863

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3884	100.0	723	22 AAB31476 Amino acid sequenc
2	3878	99.8	722	24 AAB32724 E3 ubiquitin ligas
3	3869	99.6	722	22 AAM79861 Human protein ubiq
4	3801	77.3	748	18 AAM13384 Human protein ubiq
5	3001	77.3	748	18 AAB31477 Amino acid sequenc
6	2996	77.1	804	22 AAU19610 Human diagnostic a
7	2236	57.6	514	22 AAU87301 Human central nerv
8	2069	53.3	1035	22 ABB61120 Drosophila melanog
9	1614.5	41.6	335	22 AAM78877 Human protein seq

10	1461	37.6	766	18 AAM13385
11	1436.5	37.0	832	23 AAB73459
12	1345	34.6	869	24 AAB72504
13	1318	33.9	854	20 AAY30949
14	1301	33.5	852	20 AAY30948
15	1259	32.4	927	20 AAY25170
16	1259	32.4	927	24 AAE32725
17	1245.5	32.1	739	22 ABG16477
18	1245.5	32.1	739	22 AAG68173
19	1245.5	32.1	739	22 ABR41097
20	1245.5	32.1	739	24 AAE32720
21	1245	32.1	739	24 AAE32727
22	1239	31.9	922	22 AAB50049
23	1239	31.9	922	23 ABB05708
24	1239	31.9	949	22 ABB59631
25	1239	31.9	949	22 ABB59631
26	1238.5	31.6	923	22 ABG01080
27	1228.5	31.6	995	24 AAE32719
28	1227	31.6	724	18 AAM36797
29	1226.5	31.6	759	24 AAB72504
30	1225	31.5	870	22 AAE05495
31	1225	31.5	870	24 AAE32722
32	1223.5	31.5	834	23 AAU77715
33	1223	31.5	854	24 ABB59632
34	1212	31.2	474	20 AAY10943
35	1212	31.2	506	18 AAM36795
36	1193.5	30.7	733	22 AAB50048
37	1157.5	29.8	1082	22 AAB63355
38	1151	29.6	1071	22 ABG11772
39	1150.5	29.6	818	18 AAM13386
40	1047	27.0	683	18 AAM36794
41	1047	27.0	683	22 AAE05494
42	1047	27.0	684	24 AAE32721
43	1023	26.3	380	22 AAB83047
44	1012	26.1	375	22 AAB83046
45	1002	25.8	1562	22 AAG27493

ALIGNMENTS

RESULT 1	AAAB31476 standard; Protein; 723 AA.
AC	AAB31476;
DT	20-APR-2001 (first entry)
XX	
DE	Amino acid sequence of a human Smurf1 polypeptide.
XX	
KW	Smurf1; Smurf2; Sma2 signal transduction; bone morphogenetic protein; BMP;
KW	transforming growth factor-beta; human; TGF-beta; Chondrogenesis;
KW	osteogenesis; blood differentiation; cartilage formation; hair growth;
KW	neural tube patterning; retinal development; heart induction;
KW	morphogenesis; tooth formation; gamete formation.
XX	
OS	Homo sapiens.
XX	
PN	WO200077168-A2.
XX	
PD	21-DEC-2000.
XX	
PF	12-JUN-2000; 2000MO-US16250.
XX	
PR	11-JUN-1999; 99US-0138969.
XX	
PA	(UNY) UNIV NEW YORK STATE RES FOUND.
XX	
PA	(HSCR-) HSC RES & DEV LP.
XX	
PI	Thomsen GH, Wrana J;
XX	
DR	WPI; 2001-071267/08.

Human protein ubiq
Candida albicans e
Aspergillus fumig
Murine E3 ubiquiti
Human E3 ubiquitin
Human KIAAN ligase
Human NBD4 long f
Novel human diagno
Atrophin-1 interac
Human atrophin-1 i
Atrophin-1 interac
Human ZGBP1 prote
Human homolog of D
Human signal trans
Drosophila melanog
Drosophila melanog
Novel human diagno
KIAA439 protein.
Novel human gene,
Aspergillus fumiga
Human ubiquitin pr
Nedd-4-like ubiqui
Diseased kidney ti
Human cell growth,
Amino acid sequenc
Novel human protei
Human clone 811a
Drosophila melanog
Novel human diagno
Human protein ubiq
Novel human protei
Human ubiquitin pr
Nedd-4-like ubiqui
Human WPI HECT E3
Human WPI HECT E3
Novel human diagno

DR N-PSDB; AAF24852.
 XX Novel isolated Smurf protein useful for inhibiting bone morphogenic
 PT protein or tumor growth factor-beta activation pathway, for treating
 PT cancer and to block osteogenesis, hair growth, tooth formation
 XX
 XX Claim 6; Fig 10; 107pp; English.
 XX
 CC The present sequence represents a human Smurf1 polypeptide. The
 CC specification also describes a Smurf2 polypeptide. Smurf polypeptides
 CC are negative regulators of Smad signal transduction, and antagonists of
 CC bone morphogenic protein (BMP) or transforming growth factor-beta
 CC (TGF-beta) signaling pathway. Expression of Smurf1 in a cell is useful
 CC for inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf
 CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood
 CC differentiation, cartilage formation, neural tube patterning, retinal
 CC development, heart induction and morphogenesis, hair growth, tooth
 CC formation, gamete formation and a wide variety of tissue and organ
 CC formation processes, and hinder the regeneration, growth, maintenance,
 CC etc., of bone and other tissues that are dependent on the BMP pathway.
 CC The polypeptide is useful for screening for various drugs and/or
 CC antibodies that can either enhance the BMP pathway, or inhibit it.
 CC
 XX
 XX

Sequence 723 AA;

Query Match 100.0%; Score 3884; DB 22; Length 723;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

1 GSSSIKIRLTLVCAKYLAKKDFRLLPDPFAKIVDVGSGQCHSTDVKNKTLDPKNNOHYDL 60
 1 GSSSIKIRLTLVCAKYLAKKDFRLLPDPFAKIVDVGSGQCHSTDVKNKTLDPKNNOHYDL 60
 61 YVGTDSITISVNMHKKIKKQAGAGLGVRLLSNAISRLKDTGYQRLDLCKNPSDTDA 120
 61 YVGTDSITISVNMHKKIKKQAGAGLGVRLLSNAISRLKDTGYQRLDLCKNPSDTDA 120
 121 VRGQIVVSLQTRPRITGSGVVDRCGLLENSTGYVEDSPGRPLSCFMEBAPPTDSTGA 180
 121 VRGQIVVSLQTRPRITGSGVVDRCGLLENSTGYVEDSPGRPLSCFMEBAPPTDSTGA 180
 181 AAGGNCRFVESPQDQRLQAGRLRNPVYRGSILQTPQNRPHGQSPRLPEBQRTTVOG 240
 181 AAGGNCRFVESPQDQRLQAGRLRNPVYRGSILQTPQNRPHGQSPRLPEBQRTTVOG 240
 241 QVYFLHTQVGSVTHWPRIPRLDINSVNCDELQPLPPGMEVRSYVSGRLYFVDHNNRTTQF 300
 241 QVYFLHTQVGSVTHWPRIPRLDINSVNCDELQPLPPGMEVRSYVSGRLYFVDHNNRTTQF 300
 301 TDPRLHHIMNHQCLKEPQSLPLPESGSLDEBELPAQRYERDVLQKLVRLHELSLQCP 360
 301 TDPRLHHIMNHQCLKEPQSLPLPESGSLDEBELPAQRYERDVLQKLVRLHELSLQCP 360
 361 QAGHCRIVSREBIFBESYRQIMKRPKDLKRLMYKFRGEGGLDYGVAEMLYLCH 420
 361 QAGHCRIVSREBIFBESYRQIMKRPKDLKRLMYKFRGEGGLDYGVAEMLYLCH 420
 421 MNPYVGLFOYSTNDNTYMLQINPDSSINPDHLSYFVGRIMGLAVFGHYNGGFTYF 480
 421 MNPYVGLFOYSTNDNTYMLQINPDSSINPDHLSYFVGRIMGLAVFGHYNGGFTYF 480
 481 YKOLLGKPIQLSLESVDDELAKSLVLIENDITPYLDHTFCVENNAGRLIOHELKXNG 540
 481 YKOLLGKPIQLSLESVDDELAKSLVLIENDITPYLDHTFCVENNAGRLIOHELKXNG 540
 541 RNPVTEENKKEVYRLVYVMRFRGIEAOFLALQKFNFLIQHLLKPFDDKELELLIGG 600
 541 RNPVTEENKKEVYRLVYVMRFRGIEAOFLALQKFNFLIQHLLKPFDDKELELLIGG 600
 601 LDKIDLNDKSNTRLKXCVADSNIVMFWQAVETDEERRARLLQFTVGSRTVPLOGFKA 660
 601 LDKIDLNDKSNTRLKXCVADSNIVMFWQAVETDEERRARLLQFTVGSRTVPLOGFKA 660

661 LOGTGAGPRLFTIHLIDANTDLPKATCFNRIDIPRESYEKYLEKLTAVEETCGF 720
 661 LOGTGAGPRLFTIHLIDANTDLPKATCFNRIDIPRESYEKYLEKLTAVEETCGF 720
 721 AVE 723
 721 AVE 723
 721 AVE 723

RESULT 2

AAE32724 standard; Protein; 722 AA.

AAE32724;

24-MAR-2003 (first entry)

E3 ubiquitin ligase SMURF1 protein.

Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;

poliomyelitis; HIV; measles; protein therapy; E3 ubiquitin ligase;

enzyme.

Unidentified.

WO200290549-A2.

14-NOV-2002.

12-MAR-2002; 2002MO-IB02106.

12-MAR-2001; 2001US-275224P.

31-JUL-2001; 2001US-308258P.

07-DEC-2001; 2001US-340170P.

(PROT-) PROTEOLOGICS LTD.

Greener T, Moskowitz H, Reis Y, Alroy I;

WPI; 2003-111976/10.

N-PSDB; AAD50458.

New protein complex comprising HECT-RCC1, viral maturation scaffolding

protein (VMSF), and/or HIV gag protein, useful for treating viral

infections, such as lymphosarcoma, HIV, hepatitis, poliomyelitis, measles,

or Ebola.

Disclosure; Fig 14; 150pp; English.

The invention relates to a method for modulation of viral maturation.

The invention also provides an isolated protein complex comprising a

HECT-RCC1 selected from HERT-NW, HECT-RCC1, Gag protein, Gag late

domain, p13, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2, STAM3,

VHS-TIM, GTPase, E2 enzyme, teg101, cullin, HRC1, HRC2, HRC3, Nedd4

-like protein or clathrin. The complexes, proteins, antibodies and

methods are useful for treating viral infections, such as lymphosarcoma,

human immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles,

or Ebola and for inhibiting budding in a subject. They are also useful

in diagnostic assays for determining whether a cell is infected with a

virus and for characterizing the nature, progression and/or infectivity

of the infection. The invention is also useful in protein therapy. The

```

QY 62 VGTDTSTTISVNMHKKIHKKQAGFLGCVRLSNASRLKDTGYORLDCIKNPSTDAV 121
DB 61 VGTDTSTTISVNMHKKIHKKQAGFLGCVRLSNASRLKDTGYORLDCIKNPSTDAV 120
QY 122 RGOIVVSLQTRDRIGTGGSVVDRCGLLENEGTVYEDSGGRPLSCFMEEPAVYDSTGAA 181
DB 121 RGOIVVSLQTRDRIGTGGSVVDRCGLLENEGTVYEDSGGRPLSCFMEEPAVYDSTGAA 180
QY 182 AGGNCRFVESPQDQRLQORLRNDVAGSLQTPQNRPHGQSPLEPGYQRTTVGQ 241
DB 181 AGGNCRFVESPQDQRLQORLRNDVAGSLQTPQNRPHGQSPLEPGYQRTTVGQ 240
QY 242 VYFLHTQTVSTWHDPRIPDLNSVNCDELGLPLPQMEVRSVSGRIYVDNNNTTQFT 301
DB 241 VYFLHTQTVSTWHDPRIPDLNSVNCDELGLPLPQMEVRSVSGRIYVDNNNTTQFT 300
QY 302 DPLRHHIMNHQCCQKPSQPLPLPSEGSLEDEBLPAQRYERDLVOKLKVLRHLSLQCPQ 361
DB 301 DPLRHHIMNHQCCQKPSQPLPLPSEGSLEDEBLPAQRYERDLVOKLKVLRHLSLQCPQ 360
QY 362 AGHCRIVSRREEIFEESYRQIMQRPKDLKGRIMVFRGSEGLDYGVARLWYLLCHEM 421
DB 361 AGHCRIVSRREEIFEESYRQIMQRPKDLKGRIMVFRGSEGLDYGVARLWYLLCHEM 420
QY 422 LNPYGLFQYSTDNIVYLQINPDSSINPDHLSYFHFVGRIMGLAVFHGYINGGFTVFPY 481
DB 421 LNPYGLFQYSTDNIVYLQINPDSSINPDHLSYFHFVGRIMGLAVFHGYINGGFTVFPY 480
QY 482 KQLGKPIQLSDLESVDPPELHKSIVMLTENDIPVLDHFCVEHNAFGRILOHELKPNR 541
DB 481 KQLGKPIQLSDLESVDPPELHKSIVMLTENDIPVLDHFCVEHNAFGRILOHELKPNR 540
QY 542 NVPYTEENKKEYYRLVYNWRFMGIEAQLALQXGNEELPQHLNLPDQKELELIIGGL 601
DB 541 NVPYTEENKKEYYRLVYNWRFMGIEAQLALQXGNEELPQHLNLPDQKELELIIGGL 600
QY 602 DKILNDMKSNTRLAKHCVADSNIVRWQVAVTFPDEERRARLLQFTGSTRVPLQGFRL 661
DB 601 DKILNDMKSNTRLAKHCVADSNIVRWQVAVTFPDEERRARLLQFTGSTRVPLQGFRL 660
QY 662 QGSTGAAGPRLFTIHLIDANTDNLPKAHTCFNRIDIPPYSEYKLEYEKLITAVEETGFA 721
DB 661 QGSTGAAGPRLFTIHLIDANTDNLPKAHTCFNRIDIPPYSEYKLEYEKLITAVEETGFA 720
QY 722 VE 723
DB 721 VE 722

```

```

PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSR-) HYSRQ INC.
PI Tang YT, Liu C, Dirmancic RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang Y, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR N-PSDB; AAKS2994.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 362-363; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 722 AA;

```

```

RESULT 3
AAM79861
ID AAM79861 standard; Protein; 722 AA.
XX
AC AAM79861;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3507.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
OS Homo sapiens.
XX
PN MO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001MO-US04098.
XX

```

```

QY 2 GSSIKIRLTVLCAKNAKKOFPRLPPFAKIYVDGSGGCHSTDTYKNTLDPKMOHYLY 61
DB 1 GSSIKIRLTVLCAKNAKKOFPRLPPFAKIYVDGSGGCHSTDTYKNTLDPKMOHYLY 60
QY 62 VGTDTSTTISVNMHKKIHKKQAGFLGCVRLSNASRLKDTGYORLDCIKNPSTDAV 121
DB 61 VGTDTSTTISVNMHKKIHKKQAGFLGCVRLSNASRLKDTGYORLDCIKNPSTDAV 120
QY 122 RGOIVVSLQTRDRIGTGGSVVDRCGLLENEGTVYEDSGGRPLSCFMEEPAVYDSTGAA 181
DB 121 RGOIVVSLQTRDRIGTGGSVVDRCGLLENEGTVYEDSGGRPLSCFMEEPAVYDSTGAA 180
QY 182 AGGNCRFVESPQDQRLQORLRNDVAGSLQTPQNRPHGQSPLEPGYQRTTVGQ 241
DB 181 AGGNCRFVESPQDQRLQORLRNDVAGSLQTPQNRPHGQSPLEPGYQRTTVGQ 240
QY 242 VYFLHTQTVSTWHDPRIPDLNSVNCDELGLPLPQMEVRSVSGRIYVDNNNTTQFT 301
DB 241 VYFLHTQTVSTWHDPRIPDLNSVNCDELGLPLPQMEVRSVSGRIYVDNNNTTQFT 300
QY 302 DPLRHHIMNHQCCQKPSQPLPLPSEGSLEDEBLPAQRYERDLVOKLKVLRHLSLQCPQ 361
DB 301 DPLRHHIMNHQCCQKPSQPLPLPSEGSLEDEBLPAQRYERDLVOKLKVLRHLSLQCPQ 360
QY 362 AGHCRIVSRREEIFEESYRQIMQRPKDLKGRIMVFRGSEGLDYGVARLWYLLCHEM 421
DB 361 AGHCRIVSRREEIFEESYRQIMQRPKDLKGRIMVFRGSEGLDYGVARLWYLLCHEM 420
QY 422 LNPYGLFQYSTDNIVYLQINPDSSINPDHLSYFHFVGRIMGLAVFHGYINGGFTVFPY 481

```

```

Db 421 LNPYGLFOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMGLAFHGHYINGGFTVPFY 480
Qy 482 KQLGKPIQLSDPESDPELHKSLVWLENDITPVLCHTFCVHNAPFGRITQHELKNGR 541
Db 481 KQLGKPIQLSDPESDPELHKSLVWLENDITPVLCHTFCVHNAPFGRITQHELKNGR 540
Qy 542 NVPTEENKKEYRLLVYVNRFPNGIEAGFLALQKGFNELIPQHLKPPDQKELELIIGL 601
Db 541 NVPTEENKKEYRLLVYVNRFPNGIEAGFLALQKGFNELIPQHLKPPDQKELELIIGL 600
Qy 602 DKTLDMKSNTRLEKCVADSNIVKFMQAVETPDEERRARLLQFVMSSTVPLOGFKAL 661
Db 601 DKTLDMKSNTRLEKCVADSNIVKFMQAVETPDEERRARLLQFVMSSTVPLOGFKAL 660
Qy 662 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPYSEYKLYEKLTAVEETCGFA 721
Db 661 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPYSEYKLYEKLTAVEETCGFA 720
Qy 722 VE 723
Db 721 VE 722

```

RESULT 4
AAM13384 standard; Protein; 748 AA.

```

XX AC AAM13384;
XX 10-JUL-1997 (first entry)
XX DE Human protein ubiquitin ligase publi.
XX KM Protein ubiquitin ligase; publi; cdc25 phosphatase; CDK kinase;
XX KM p53; cell cycle; transgenic animal.
XX OS Homo sapiens.
XX PN M09712962-A1.
XX PD 10-APR-1997.
XX PF 04-OCT-1996; 96WO-US15930.
XX PR 04-OCT-1995; 95US-0539205.
XX PA (COLD-) COLD SPRING HARBOR LAB.
XX PT Beach D, Caligiuri M, Nefsky B;
XX DR WPI; 1997-226206/20.
XX DR N-PSDB; AAT47040.
XX PT Human and Saccharomyces pombe protein ubiquitin ligase (e) - involved
XX PT in cdc25 phosphatase and p53 ubiquitination, and regulate cell
XX PT growth and proliferation
XX PS Claim 1; Page 74-77; 108pp; English.

```

Human protein ubiquitin ligases publi (AAM13384), pub2 (AAM13385) and pub3 (AAM13386) are homologues of fission yeast publi (AAM13387) and were identified from cDNA clones (AAT47040-42) obtcd. e.g. from a keratinocyte cDNA library. Pub polypeptides can be produced in transfected host cells. They can control the steady state level of cdc25 phosphatase, the degree of CDK kinase (e.g. cdc2) dephosphorylation and the steady state level of p53 (controlling the degree of cell cycle regulation of p53). They can regulate cell or tissue differentiation, or cell growth or proliferation by affecting other proteins, can be a specific (ant)agonist of wild-type protein function and may be used as immunogens to elicit a specific immune residue.

```

SQ Sequence 748 AA;
Query Match 77.3%; Score 3001; DB 18; Length 748;
Best Local Similarity 74.6%; Pred. No. 6.4e-286;
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

Qy 5 IKIRLVLCANLAKKOPFRLDPDPAKIVVDSGGCHSTDTVXNLTDPKKNQHYDLYVK 64
Db 12 VILRLTVLCANLAKKOPFRLDPDPFAVVDGSGCHSTDTVXNLTDPKKNQHYDLYVK 71
Qy 65 TDSITISVWNNKKIHKKQAGFLGCVLLSNATSRLLKDTGYRLDCKNPSPTDAVAGQ 124
Db 72 SDSVTISVWNNKKIHKKQAGFLGCVLLSNATSRLLKDTGYRLDCKNPSPTDAVAGQ 131
Qy 125 IVLSLQTRDRIIGTGGVVDGRLLENL-----GTYY----- 155
Db 132 IVLSLQSRDRIIGTGGVVDGRLLENLDPGWEERRATSGRIQYLNHTTPTTQWERPTSP 191
Qy 156 --EDSGRPRPLSCMEERFAPYTDSTGAAGGNRPFVESPQDORLQACRLRPDVRGSL 213
Db 192 ASEYSPGRPLSCFVDENTPISGTGATCG-----OSSDPLERRVRQRHNTAT 242
Qy 214 QTPQRRPHGSPFLPEGEBOQTVOGVFLHTQGVSTWHDPRILPRLDINSVNCDELGP 273
Db 243 ---SRTHLTPPDLPEGEBOQTVOGVFLHTQGVSTWHDPRVPRDLNINCEELGP 298
Qy 274 LPPGWEVASTVSGRIYVDHNRTTQTTDR---LHINNHCQCKEPGQPLPLSEGS 329
Db 299 LPPGWEINNTATGAVYFVDHNNRTTQTTDRPLSANLHVLNRRONQLDQOOQOV---S 354
Qy 330 L---EDEELPAQRYERDLYOCLKTLRHELSLQOPQAGCHRIEVSREIEPESYRQIMKOR 386
Db 355 LCPDTECLTVPRKRDLYOCLKTLRQELSQOQAGCHRIEVSREIEPESYRQIMKOR 414
Qy 387 PRDLKKRLMKFRGEBGLDYGVARREYLILCHEMLNPYGLQYSTDNITVMIQINPDS 446
Db 415 PRDLKKRLMKFRGEBGLDYGVARREYLILSHEMLNPYGLQYSTDDIYTLQINPDSA 474
Qy 447 INPDLSYFHFVGRIMGLAFHGHYINGGFTVPFYKQLGKPIQLSDPESDPELHKSLV 506
Db 475 VNPEHLSYFHFVGRIMGLAFHGHYINGGFTLPFYKQLGKSLTLDMDLVDLDNLSLY 534
Qy 507 WLENDITPVLCHTFCVHNAPFGRITQHELKPNGRNVPTEENKKEYRLLVYVNRFPNGI 566
Db 535 WLENDITGVLDHTEFCVHNAPGEIILQHELKPNKGSIPVNEENKKEYRLLVYVNRFPNGI 594
Qy 567 EAQFLALQKGFNELIPQHLKPEPOCKELELIIGSLDKIDIDMKSNTRLEKHCYVADSNIYK 626
Db 595 EAQFLALQKGFNEVLPQHLKTFPEKELELIIGSLGKIDIDMKSNTRLEKHCYVADSNIYK 654
Qy 627 WFMQAVETPDEERRARLLQFVGTSTRVPLQGFALQGSTGAAGPRLFTIHLIDANTDNL 686
Db 655 WFMQAVETPDEERRARLLQFVGTSSRVPLQGFALQ---GAAGPRLFTIHLIDANTDNL 711
Qy 687 KATCFNRIDIPYSEYKLYEKLTAVEETCGFAVE 723
Db 712 KATCFNRIDIPYSEYKLYEKLTAVEETCGFAVE 748

```

RESULT 5
AAB31477 standard; Protein; 748 AA.
AAB31477;
AAB31477;
20-APR-2001 (first entry)
Amino acid sequence of a human Smurf2 polypeptide.
Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP; transforming growth factor-beta; human; TGF-beta; chondrogenesis; osteogenesis; blood differentiation; cartilage formation; hair growth; neural tube patterning; retinal development; heart induction;

KM morphogenesis; tooth formation; gamete formation.
 XX Homo sapiens.
 OS
 XX MO200077168-A2.
 PN
 XX 21-DEC-2000.
 PD
 XX 12-JUN-2000; 2000WO-US16250.
 PF
 XX 11-JUN-1999; 99US-0138969.
 PR
 XX (UNYV) UNIV NEW YORK STATE RES FOUND.
 PA (HSCR-) HSC RES & DEV LP.
 XX
 PI Thomsen GH, Wraana J;
 XX
 DR MPI, 2001-071267/08.
 DR N-PSDB; AAF24853.
 XX
 XX Novel isolated Smurf protein useful for inhibiting bone morphogenic
 PT protein or tumor growth factor-beta activation pathway, for treating
 XX cancer and to block osteogenesis, hair growth, tooth formation
 PS Claim 10; Fig 12; 107bp; English.

XX The present sequence represents a human Smurf2 polypeptide. The
 CC specification also describes a Smurf1 polypeptide. Smurf polypeptides
 CC are negative regulators of Smad signal transduction, and antagonists of
 CC bone morphogenic protein (BMP) or transforming growth factor-beta
 CC (TGF-beta) signaling pathway. Expression of Smurf1 in a cell is useful
 CC for inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf
 CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood
 CC differentiation, cartilage formation, neural tube patterning, retinal
 CC development, heart induction and morphogenesis, hair growth, tooth
 CC formation, gamete formation and a wide variety of tissue and organ
 CC formation processes, and hinder the regeneration, growth, maintenance,
 CC etc., of bone and other tissues that are dependent on the BMP pathway.
 CC The polypeptide is useful for screening for various drugs and/or
 CC antibodies that can either enhance the BMP pathway, or inhibit it.

XX Sequence 748 AA;

Query Match 77.3%; Score 3001; DB 22; Length 748;
 Best Local Similarity 74.6%; Pred No. 6.4e-286;
 Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

QY 5 IKRLITVLCANLAKKDPFRLLPDPFAKIVVDSGGQCHSTDTVKNLTDPKMNQHYDLYGK 64
 DB 12 VKRLITVLCANLAKKDPFRLLPDPFAKIVVDSGGQCHSTDTVKNLTDPKMNQHYDLYGK 71
 QY 65 TDSITISVWNHKKIHKKQAGFLGCVRLISNAISRLKDTGVORLDICTLNSDPAVGRQ 124
 DB 72 SDSVTISVWNHKKIHKKQAGFLGCVRLISNAISRLKDTGVORLDICTLNSDPAVGRQ 131
 QY 125 IVVSLQTRDRIGTSGVDCRGLLENE-----GVVY----- 155
 DB 132 IVVSLQTRDRIGTSGVDCRGLLENE-----GVVY----- 155
 QY 156 --EDSGRGRPLSGFMEEPAFYDSTGAAGGNCVFESPSDQRLQARLNDVGRSL 213
 DB 192 ASEYSSPGRPLSGFVDENTPIGTINGATCG-----QSSDPRLLARRSRHRNYM 242
 QY 214 QTPQNRPHGOSPELPEGEQRTTVQGVYFLATDTGVSTWHPRIIPDLNSVNCDELGP 273
 DB 243 ---SRTHLTPDLEGEQRTTVQGVYFLATDTGVSTWHPRIIPDLNSVNCDELGP 298
 QY 274 LPPGMEVRSVSGRIYFVHNNTTFTDPR---LHIANHQCCKSPSPPLPSEBGS 329
 DB 299 LPPGMEVRSVSGRIYFVHNNTTFTDPR---LHIANHQCCKSPSPPLPSEBGS 354
 QY 330 L--EDELPAQYEDLVOKLVLEHLSLOPOAGHCRLEVSREELFEESYQIMKR 386

DB 355 LCPDTECLTVPRYKRDLYOKLILNQLSQQCPQAGHCRLEVSREELFEESYQIMKR 414
 QY 367 PKDKKRLMVFREGEGLDYGVARRELYLLGCEMNPNPYGLFQYSTDNITVMQINPDS 446
 DB 415 PKDKKRLMVFREGEGLDYGVARRELYLLGCEMNPNPYGLFQYSTDNITVMQINPDS 474
 QY 447 INPDLSYFHFVGRIMGLAVFHGYINGGFVPRYQKLGKPIQLSDLESVPDLHLSIV 506
 DB 475 VNPEHLSYFHFVGRIMGLAVFHGYINGGFVPRYQKLGKPIQLSDLESVPDLHLSIV 534
 QY 507 WLENDITPVLDHFFCVENHAPFGIILCHELKPGRNVPVTEENKKEVYVLYNWRFRMGI 566
 DB 535 WLENDITPVLDHFFCVENHAPFGIILCHELKPGRNVPVTEENKKEVYVLYNWRFRMGI 594
 QY 567 EAOFLALQKGFNEIIPQHLKPPDOKELIELIGLDPKIDLMDKSNTRLKCVASNIYR 626
 DB 595 EAOFLALQKGFNEIIPQHLKPPDOKELIELIGLDPKIDLMDKSNTRLKCVASNIYR 654
 QY 627 WFWQAVTFDEERARLLOPTGSTRVFLQGFALQGSTAAGPRFLTTHLDANTDNL 686
 DB 655 WFWQAVTFDEERARLLOPTGSTRVFLQGFALQGSTAAGPRFLTTHLDANTDNL 711
 QY 667 KAHTCFNRIDIPYSEYEKLYEKLTAVEETCGFAVE 723
 DB 712 KAHTCFNRIDIPYSEYEKLYEKLTAVEETCGFAVE 748

RESULT 6
 AAU19610
 ID AAU19610 standard; Protein; 804 AA.
 XX
 XX AAU19610;
 AC 04-DEC-2001 (first entry)
 XX
 XX Human diagnostic and therapeutic polypeptide (DITHP) #196.
 XX
 XX Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
 XX cell proliferative disorder; Crohn's disease; lymphoma; leukemia;
 XX acquired immune deficiency syndrome; AIDS; autoimmune disorder;
 XX respiratory disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200169927-A2.
 PN
 XX 30-AUG-2001.
 PD
 XX 21-FEB-2001; 2001WO-US06059.
 XX
 XX 24-FEB-2000; 2000US-0184693.
 PR 24-FEB-2000; 2000US-0184697.
 XX 24-FEB-2000; 2000US-0184698.
 PR 24-FEB-2000; 2000US-0184768.
 XX 24-FEB-2000; 2000US-0184769.
 PR 24-FEB-2000; 2000US-0184770.
 XX 24-FEB-2000; 2000US-0184771.
 PR 24-FEB-2000; 2000US-0184772.
 XX 24-FEB-2000; 2000US-0184773.
 PR 24-FEB-2000; 2000US-0184774.
 XX 24-FEB-2000; 2000US-0184775.
 PR 24-FEB-2000; 2000US-0184776.
 XX 24-FEB-2000; 2000US-0184777.
 PR 24-FEB-2000; 2000US-0184797.
 XX 24-FEB-2000; 2000US-0184813.
 PR 24-FEB-2000; 2000US-0184817.
 XX 24-FEB-2000; 2000US-0184841.
 PR 24-FEB-2000; 2000US-0185213.
 XX 24-FEB-2000; 2000US-0185215.
 PR 12-MAY-2000; 2000US-0203785.
 XX 15-MAY-2000; 2000US-0204525.
 PR 16-MAY-2000; 2000US-0204525.
 XX 16-MAY-2000; 2000US-0204821.
 PR 16-MAY-2000; 2000US-0204908.

PR 16-MAY-2000; 2000US-0205232.
 PR 17-MAY-2000; 2000US-0204815.
 PR 17-MAY-2000; 2000US-0204863.
 PR 17-MAY-2000; 2000US-0205221.
 PR 17-MAY-2000; 2000US-0205285.
 PR 17-MAY-2000; 2000US-0205286.
 PR 17-MAY-2000; 2000US-0205287.
 PR 17-MAY-2000; 2000US-0205323.
 PR 17-MAY-2000; 2000US-0205324.
 PA (INCY-) INCYTE GENOMICS INC.
 XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup NS, Chang SC;
 PI Chen A, D'Sa SA, Amstey S, Dahl CR, Dam TC, Daniels SE,
 PI Dufour GE, Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL,
 PI Liu TF, Roseberry AM, Russo BH, Russo FD, Stockbrener TK, Daffo A;
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W,
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;
 XX WPI; 2001-502867/55.
 DR N-PSDB; AAS31181.
 XX
 PT Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics
 PT
 PS Claim 27; Page 512-514; 522pp; English.
 XX
 CC The invention relates to polynucleotides (I) encoding diagnostic and
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes,
 CC and proteins involved in growth and development and receptors. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate DITHP expression. For example, (I) and
 CC (II) may be used to treat disorders associated with decreased polypeptide
 CC expression by rectifying mutations or deletions in a patient's genome,
 CC that affect the activity of the DITHPs, by expressing inactive proteins
 CC or supplementing the patient's own production of them. (I) and (II)
 CC may be used to treat diseases, for example, cell proliferative disorder,
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. (I) and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. (II) may also be used as antigens in the production of
 CC antibodies against DITHPs and in assays to identify modulators of DITHP
 CC expression and activity. The anti-DITHP antibodies and antagonists may
 CC also be used to down regulate expression and activity. The anti-DITHP
 CC antibodies may also be used as diagnostic agents for detecting the
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbant
 CC assay (ELISA)). AAU19415-AAU19625 represent human diagnostic and
 CC therapeutic (DITHP) polypeptides of the invention.
 XX
 SQ Sequence 804 AA;
 Query Match 77.1%; Score 2996; DB 22; Length 804;
 Best Local Similarity 74.5%; Pred. No. 2.3e-285;
 Matches 564; Conservative 63; Mismatches 72; Indels 58; Gaps 8;
 QY 5 IKIRLTVCANRKLAKKQDFRLPDPFAKIYVDGSGQSGSTVTYVKNTPDPKXNOMYDLYVKK 64
 DB 68 VKHLTLTVCAKRLVKQDFRLLPDPFAKVVVDGSGQSGSTVTYVKNTPDPKXNOMYDLYVKK 127
 QY 65 TDSITISVWNHKKTHKKQAGAGFLCCVRLSNALSRKDTGYORLDCNLPSTDAVRGQ 124
 DB 128 SDSVTISVWNHKKTHKKQAGAGFLCCVRLSNALSRKDTGYORLDCNLPSTDAVRGQ 187
 QY 125 IVVSLQTRDRIGTGGSVYDRCGLTENE-----GTVV----- 155
 DB 188 IVVSLQTRDRIGTGGSVYDRCGLTENE-----GTVV----- 247
 QY 156 --EDSGPGRPLSCWEPAPYTDSTGAAGGCGNCRFVESPSQDRLQAOGLRNPVGRSL 213

DB 248 ASRYSEPGPLSCFPDENTPISTGACG-----OSSDRLERVRSGRRHNYM 298
 QY 214 QTPONRPHGQSPPELPEGEORRTVQGVYFLATOTGYSWHDRIIPRDNVNCDELGP 273
 DB 299 ----SKTHLHTPDPPEGEORRTVQGVYFLATOTGYSWHDRIIPRDNVNCDELGP 354
 QY 274 LPPGWEKSTVSGRIYFVDHNNRTTQFTDPR-----LHHNNHQQQLKEPQPLPPEGS 329
 DB 355 LPPGWEKSTVSGRIYFVDHNNRTTQFTDPR-----LHHNNHQQQLKEPQPLPPEGS 410
 QY 330 L---EDDELPARQYERDLYQKXLRHLSLQOPAGCRTEVERRETFEESYQIMKR 386
 DB 411 LCPDDTECLTVPRYKRDLYQKXLRHLSLQOPAGCRTEVERRETFEESYQIMKR 470
 QY 387 PKDLKRLMVKRGEGLDYGVAREMVLKHEMLNYPGLFOYSTDNIMYLQINDPS 446
 DB 471 PKDLKRLMVKRGEGLDYGVAREMVLKHEMLNYPGLFOYSTDNIMYLQINDPS 530
 QY 447 INPDHSTYHFGRIWGLAVFHHGHTINGFTVPPKOLLGPQLSDLESVDELHKSIV 506
 DB 531 VNEHLSYHFGRIWGLAVFHHGHTINGFTVPPKOLLGPQLSDLESVDELHKSIV 590
 QY 507 WLENDITFVLDTFCVEHNAFGRILOHELKPNGRNVPTEENKKEYRLVYNNRFNGI 566
 DB 591 WLENDITFVLDTFCVEHNAFGRILOHELKPNGRNVPTEENKKEYRLVYNNRFNGI 650
 QY 567 EAQFLALQGFNEVLPQHLKFPDQKEELTIGLKDILDMKSNTRKQCVADSNIVR 626
 DB 651 EAQFLALQGFNEVLPQHLKFPDQKEELTIGLKDILDMKSNTRKQCVADSNIVR 710
 QY 627 WFOAVETDEERRARLLOFVTSRTRVPLQGFKALQSGTGAAPRLFTIHLIDANTDLP 686
 DB 711 WFOAVETDEERRARLLOFVTSRTRVPLQGFKALQSGTGAAPRLFTIHLIDANTDLP 767
 QY 687 KATCFNRIDIPYBSYKCLYKLTAVETGFAVE 723
 DB 768 KATCFNRIDIPYBSYKCLYKLTAVETGFAVE 804
 RESULT 7
 ID AAU87301 standard; Protein; 514 AA.
 XX AAU87301;
 AC
 AC
 DT 05-JUN-2002 (first entry)
 XX
 DE Novel central nervous system protein #211.
 XX
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasia; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; ischemia; angioneurosis;
 KW nervous system disorder; Alzheimer's disease; ocular disorder;
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
 KW adenocarcinoma; reproductive system disorder; testicular feminization;
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularization;
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;
 KW myocardial infarction; wound healing; cell proliferation; skin aging;
 KW food preservative; gene therapy.
 OS Homo sapiens.
 XX
 XX W0200155318-A2.
 PD 02-AUG-2001.
 XX
 PD 17-JAN-2001; 2001WO-US01332.
 PF
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220966.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225759.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226281.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227189.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228992.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0228344.
 PR 01-SEP-2000; 2000US-0228345.
 PR 05-SEP-2000; 2000US-0228309.
 PR 05-SEP-2000; 2000US-0228513.
 PR 05-SEP-2000; 2000US-0228513.
 PR 06-SEP-2000; 2000US-0230438.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0233400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 01-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251866.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-581633/65.
 N-PsDB; ABK43631.
 New isolated nucleic acid encoding a protein for diagnosing,
 preventing, treating or ameliorating medical conditions and used as
 food additives or preservatives -
 Claim 9; SEQ ID No 819; 837bp; English.

CC The invention describes an isolated nucleic acid molecule (I) encoding a
 CC novel central nervous system protein. (II) and polypeptides (III) encoded
 CC by (I), are used to treat a medical conditions and in diagnosis of a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. Rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
 CC adenocarcinomas and irritable bowel syndrome, reproductive system
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
 CC leukemia, disorders involving neovascularisation e.g. melaninactes,
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
 CC acute kidney failure and blood related disorders e.g. myocardial
 CC infarction. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 57.6%; Score 2236; DB 22; Length 514;
 Best Local Similarity 80.7%; Pred. No. 8.8e-211;
 Matches 414; Conservative 43; Mismatches 42; Indels 14; Gaps 4;

QY 218 NRPHGOSPELPGYVORTVGVGVFHTQGVSTWHPRIPLDINSVDCDEGLPLPG 277
 DB 9 SRTHHTPPDLPGYVORTVGVGVFHTQGVSTWHPRIPLDINSVDCDEGLPLPG 68
 QY 278 WEVSVSGRIYVVDNNRTTFTDPR---LHIMNHOCQKESQPLPSPSEGL-- 330
 DB 69 WEIRNTATGRVYFVDHNNRTTFTDPRLSANHLVLRNQKLDQOQOQV---SLCPD 124
 QY 331 EDELPAGQRYERPLVOKLYLREHLSLOQPOAGHCHIEVSEIFEESYROMKWRPDL 390
 DB 125 DTECLTVPRYKRDVQKTLRLSELQOQPOHCHIEVSEIFEESYROMKWRPDL 184
 QY 391 KKRIMVFRGEGLDYGVAWEVLYLHEMLNPYYGLFOYSTDNITMLOINPDSSIND 450
 DB 185 WKRLMKFRGEGLDYGVAWEVLYLHEMLNPYYGLFOYSTDNITMLOINPDSSIND 244
 QY 451 HLSYFHVGVIMGLAVHGHYINGGFTVPPYKOLGKPIQLSDLESVPBELKSLWTLE 510
 DB 245 HLSYFHVGVIMGLAVHGHYINGGFTVPPYKOLGKSLTLDMEVDPDLNLSWTLLE 304
 QY 511 NDITPVLDRHFCVEHNAFGRILCHELKEFGNRPVTEENKKEVRLYVWRFMRGIEAQF 570
 DB 305 NDITGVLDHFCVEHNAFGEIICHELKPNKGSIPVNEENKEVRLYVWRFMRGIEAQF 364
 QY 571 LALOKGFNELIPQHLKPPQKLELILIGLDKIDLNPKSNTRLKCYADSNITRWQ 630
 DB 365 LALOKGFNEVIPHLLKTFDEKELELILIGLDKIDVDMKVNTRLHCPDSDNITRWQ 424
 QY 631 AVETFEDEERRARLLQFTVGSSTRVPLQGFALQSGSTAGAPRLFTIHLIDANTDLPKAT 690
 DB 425 AVETFEDEERRARLLQFTVGSSTRVPLQGFALQ---GAAPRLFTIHOIDACNNPKAT 481
 QY 691 CFNRIDIPYESTIEKLYEKLITAVEETCGFAVE 723
 DB 482 CFNRIDIPYESTIEKLYEKLITAVEETCGFAVE 514

RESULT 8
 ABB61120
 ID ABB61120 standard; Protein; 1035 AA.
 AC ABB61120;
 XX
 XX
 DT 26-MAR-2002 (first entry)
 XX

DE Drosophila melanogaster polypeptide SEQ ID NO 10152.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN MO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW,

XX WPI; 2001-656860/75.

DR N-PSDB; ABL05223.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PS interactions -

PS Disclosure; SEQ ID NO 10152; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB16176-AB120511), expressed DNA

XX sequences (AB101840-AB116175) and the encoded proteins

XX (AB857373-AB872072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 1035 AA;

QY Query Match 53.3%; Score 2069; DB 22; Length 1035;

QY Best Local Similarity 41.2%; Pred. No. 8.6e-194;

QY Matches 435; Conservative 104; Mismatches 145; Indels 372; Gaps 15;

QY 6 KIRLTVLCAKNIKAKDFRLDPDPFAKIVVDGSGCHSTIDVYKNTLDPRKMQHDIYVGT 65

QY 14 KVRITTLCAKNIKAKDFRLDPDPFAKIVVDGSGCHSTIDVYKNTLDPRKMQHDIYVGT 73

QY 66 DSITISVNHKKIKHKKQAGFAGCYRLSNASIRLKDTGYORLDCGLNPSDIDAVAGQI 125

QY 74 DATITVVMQRKIKH--GSGFLGCVRIAPFNIQSLKAGFRLDLGLSPDDDLVAGQI 131

QY 126 VVSLQTRDRIGTG---GSAVDCRGLLENE-----GTVY----- 155

QY 132 IISLSKDPSSGPNPLATVPGSDVGRGSEDDSSDSLSPEGSEERRDNGRVVYVNHATK 191

QY 156 ----- 155

QY 192 STQWDRPRPGVVGSSGSHATSPQQRNTNNGSGDRQAPAGTRSTTCTNLMNNHRSRL 251

QY 156 -----EDSGRPLSCFMEBPADYDSTAAAGGACRFVE----- 191

QY 252 SVTASDERHSTELISVYKENTSPPTVSA--TTTPCKTSSSSSSSGG--RTLEORPT 308

QY 192 ----SPSQQRLOAQRINPDVRSGLQPRPFGHOSPE----- 227

QY 309 NEPATPTSTTSASVRLHNSD--NHVKIPKQCTGNHAPPESTPTSPGCGQNVVGNAGNG 366

QY 228 ----- 227

Db	367	STSGNSGGAQAPQASNSGWTQEDAAITTSPTTSSPPRHQSFPPTNIPSPASVTPSAN	426
Qy	228	-----	227
Db	427	GNVHSEFNANSTPAGSGGGSRYTAATPGQSRGSRSSQGEESTRRRSKGTNGTSG	466
Qy	228	-----LPEVEERTVVOQVYFLHTQGVSTWHDPR	259
Db	487	GGGGGGSGGRYASAAIAAANAAPPFLDPPEGEYENRTQQGQVYFTHLPTGVSTWHDPR	546
Qy	260	PRDLNS--VNCDELGPDPGMEVSTVSGRIYFVHNHNRRTQFTDPRLHHIM-----	309
Db	547	PRPDFDQHLLDAIGLPSGMEGRKASGRVYFVDHNHNRRTQFTDPRASGSLQMIKRG	606
Qy	310	---NHQCQKZBSPQPLPS-----EG	328
Db	607	VPTPSAANACTPAPSPATPAPSAABAABVPOATPASNATPTTLTTTTPPHRIVPDLPOG	666
Qy	329	SLEDEELPAQRYERDVLQKTKVLREHLSLQOPQAGHCLEVSREEIFPESYRQIMKRPK	368
Db	667	LLEBADL-LPKYKRDVLGKLRALRTLTQTMQDSGCHLEVSRENEIFPESYRLIMKRAK	725
Qy	389	DLKRLMWKXRGEEGLDYGVAWEMLYLLCHEMLNPYGLFOYSTDNITYMLQINPDSSIN	448
Db	726	DMRKRLMWKFKGEEGLDYGVAWEMLHLISREMLNPQGLFOYSRDDHYTLQINPDSGVN	785
Qy	449	PDHLSTYHFLRIMGJLAVFHGHYINGFLVPPYKQLLQKPIQLSLSEVDDELKSLVMI	508
Db	786	PDHLSTYHFL-----LNNPITLGDIBGVDDPLHSLSTM	819
Qy	509	LENDITPVLDHTECVENHAFGRLLQHELKPNRNVPTYEENKEEYVRLVYVMRPMRGIEA	566
Db	820	LENSISGLIESTSVENNSFGALVYHELLDPGASIPVTEENKREYVKLYVYRPMRGIEQ	879
Qy	569	QFLALQGFENELLPOHLKPFQDKXELLLIGLIDKIDLNDWKSNTRLKGVADSNIYRMF	628
Db	880	QFLALQGFPELLPSHLRPFDERELELVIGIGISIDVNDMNNTRLKGCTNETQVLMWF	935
Qy	629	WQAVETFEDEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGRLFTIHL-IDANTDMLPK	687
Db	940	WQAVETSSSEMRARLLQFVTGSTRVPLQGFKALQSGTGAAGRLFTIHLTDVLPQMLPK	995
Qy	688	AHTCFNRIDIPYSESYEKLYEKLLTAAVEETCGFAVE	723
Db	1000	AHTCFNRIDIPYETVYQLLQCDKLTQAVEETCGFAVE	1035
RESULT 9			
AAAT78877			
ID	AAAT78877	standard; Protein; 335 AA.	
XX	AAAT78877;		
XX	AC		
XX	AT	06-NOV-2001 (first entry)	
XX	XX		
DE	XX	Human protein SEQ ID NO 1539.	
XX	XX		
XX	XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KV	XX	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KV	XX	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KV	XX	nervous system disorder; arthritis; inflammation.	
OS	XX	Homo sapiens.	
XX	XX		
XX	XX	WQ200157190-A2.	
XX	XX		
PD	XX	09-AUG-2001.	
XX	XX		
PE	XX	05-FEB-2001; 2001MO-US04098.	
XX	XX		
XX	XX	03-FEB-2000; 2000US-0496914.	
PR	XX	27-APR-2000; 2000US-0560875.	
PR	XX	20-JUN-2000; 2000US-0598075.	

PR	19-JUL-2000;	2000US--0620325.	
PR	01-SEP-2000;	2000US--0654935.	
PR	15-SEP-2000;	2000US--0663551.	
PR	20-OCT-2000;	2000US--0693325.	
PR	30-NOV-2000;	2000US--0728422.	
PA	(HYSE-) HYSEQ INC.		
XX			
XX	Tang YT, Liu C, Drmanac RT, Aoudi V, Zhou F, Xu C, Cao Y, Ma Y;		
FI	Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;		
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;		
XX			
DR	WPI: 2001-476283/51.		
DR	N-PSDB; AAK52010.		
XX			
FT	Nucleic acids encoding polypeptides with cytokine-like activities,		
PT	useful in diagnosis and gene therapy -		
XX			
PS	Claim 20; Page 3847; 6221p; English.		
XX			
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the		
CC	encoded polypeptides (AAM78323-AAK89302) that exhibit activity elating to		
CC	cytokine, cell proliferation or cell differentiation of which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	e.g. stem cell growth factor activity, haematopoiesis regulating		
CC	activity, tissue growth factor activity, immunomodulatory activity and		
CC	activin/inhibin activity and may be useful in the diagnosis and/or		
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and		
CC	inflammation.		
CC	Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666		
CC	(AAM80020) are omitted as the relevant pages from the sequence listing		
XX	were missing at the time of publication.		
XX			
SO	Sequence 335 AA;		
XX			
Query Match	41.6%; Score 1614.5; DB 22; Length 335;		
Best Local Similarity	86.2%; Pred. No. 7.8e-150;		
Matches 306; Conservative	4; Mismatches 20; Indels 25; Gaps 2;		
QY	168 MERRPARYTSTGAAGGNCRFESPSODQLQRLRPNDVRSGLQTPONRPHGQSE	227	
DB	1 MERRPARYTSTGAAGGNCRFESPSODQLQRLRPNDVRSGLQTPONRPHGQSE	60	
QY	228 LPEGYQRFTTGGQVYFLHTQTGVSWMHPRIIPRLDINSVACDELGLPPGWEVRSTVSGR	287	
DB	61 LPEGYQRFTTGGQVYFLHTQTGVSWMHPRIIPRLDINSVACDELGLPPGWEVRSTVSGR	120	
QY	288 IYVDVHNNTTFTPTDRLHIMNHQCCQKESQPLPLPSGSLADELPAQRYERDLYQK	347	
DB	121 IYVDVHNNTTFTPTDRLHIMNHQCCQKESQPLPLPSGSLADELPAQRYERDLYQK	180	
QY	348 LKYLRLRLSLQOPQPGHRIEVSREIPEESRYRIMQKDKKILMKRFGEGLDYG	407	
DB	181 LKYLRLRLSLQOPQPGHRIEVSREIPEESRYRIMQKDKKILMKRFGEGLDYG	240	
QY	408 GVAREMLYLILCHEMLNPPYGLFQYSTDNITVYLQINPDSINPDHLSYFHFVGRINGLAVF	467	
DB	241 GVAREMLYLILCHEMLNPPYGLFQYSTDNITVYLQINPDSINPDHLSYFHFVGRINGLAVF	298	
QY	468 HGVYINGSTVTFYFKQLGKPIQJSDLESVDPELHKSILVILENDITPVLDHTFC	522	
DB	299 -----CVPEWTLHQRG-----LHSAILDQAAGGAHPALRSGLC	330	
RESULT 10			
AAM13385			
ID	AAM13385 standard; Protein: 766 AA.		
AC	AAM13385;		
XX			
DT	10-JUL-1997 (first entry)		

XX DE Human protein ubiquitin ligase pub2.
 XX KM Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase;
 XX KM p53; cell cycle; transgenic animal.
 XX OS Homo sapiens.
 XX PN MO9712962-AL.
 XX PD 10-APR-1997.
 XX PF 04-OCT-1996; 96MO-US15930.
 XX PR 04-OCT-1995; 95US-0539205.
 XX PA (COLD-) COLD SPRING HARBOR LAB.
 XX PI Beach D, Caligiuri M, Nefsky B;
 XX DR WPI; 1997-226206/20.
 XX DR N-PSDB; AAT47041.
 XX PT Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved
 PT in cdc25 phosphatase and p53 ubiquitination, and regulate cell
 PT growth and proliferation
 XX PS Claim 1; Page 80-84; 108pp; English.
 XX CC Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and
 CC pub3 (AAW13386) are homologues of fission yeast pub1 (AAW13387) and
 CC were identified from cDNA clones (AAT47040-42) obt'd. e.g. from a
 CC keratinocyte cDNA library. Pub polypeptides can be produced in
 CC transfected host cells. They can control the steady state level of
 CC cdc25 phosphatase, the degree of CDK kinase (e.g. cdc2)
 CC dephosphorylation and the steady state level of p53 (controlling
 CC the degree of cell cycle regulation of p53). They can regulate
 CC cell or tissue differentiation, or cell growth or proliferation by
 CC affecting other proteins, can be a specific (ant)agonist of wild-
 CC type protein function and may be used as immunogens to elicit a
 CC specific immune residue.
 XX SQ Sequence 766 AA;
 Query Match 37.6%; Score 1461; DB 18; Length 766;
 Best Local Similarity 40.8%; Pred. No. 4.5e-134;
 Matches 326; Conservative 118; Mismatches 236; Indels 120; Gaps 20;
 QY 4 SIKRLTVLCAKNAKKDPFRLPDPFAKIYVDSGQCHSTDTYKNTLDPKMGHYDLYVG 63
 DB 7 SRIKIVIVAAADGKYKRDVFRFPDPFAVLTVDSG-QHTTTAIKKTLLPYMNETFEVNT 65
 QY 64 KTDSTTISVMNHKKIKHKGAGPLGCVRLSNALSKDGYQL--DLCKNPSPTDLY 121
 DB 66 DNSTAIQVPOCKE-KKGGQGLGVINLRVGDVLDLIGDEMLICDLKLN--ENTTV 122
 QY 122 RQGIIVSLQ-----TRDRIGTGSVYDCKGLLENGETYED 157
 DB 123 HGKIIINLSTAGQLTVPSAASGARTORTSINDPOSSKSSVSNNPSSAGSPTRD 182
 QY 158 SGPG-----RPLSCFMEZ---PAPY---TSTG-----AAA 182
 DB 183 NAPAASPASSEPRFTSSFDQGRIPGEMRCITNINLRITVVDHNTKSTWIRPNLSVA 242
 QY 183 GGGNCRFVESPS-----ODRLQARLNPVVRGSLQTPQNRPHGOSPELPFGVYQ 234
 DB 243 GAAAEHLSSASNAVTEGVQPSNNARTEAVALSNMTTNG-----SGELLGWQ 296
 QY 235 RTTVGGVYFLHTQGVSTWDPRIPRDLNSV-----CDELGLPFGMEVST 283
 DB 297 RYTPGGRYFVDHNRRTTWDPRRQYIRSYGGPNNAITQQQPFVSGLPFGMEWRLLT 356
 QY 284 VSGRIYFVDHNRRTQFTDPRLLHIMNQCOLKPSQPLPSESGLEDELPQGRYERD 343

DB 357 NTARVYFVDHNRKTTTWDPR-----LPS--SL-DQVVF--QYKRD 392
 QY 344 LVQCKVTLRHELSLQOPQAGCRIVSRSEIFPESYQIMMRKDLKRLMVFREDEG 403
 DB 393 FRKKLYFLSQPAL-HPLPGQCHIVRNHLPFESYALIMQSATDLKRLMTFDSBG 451
 QY 404 LDYGVAREMLYLICHEMLNPFYGLFOYSTDNIVMLQINPDSINPHLSYHFVGRIMG 463
 DB 452 LDYGSLSREYFFLLSHENMNPYCLFEISSVDNTLQINPDSINPHEINFKIGVIG 511
 QY 464 LAVFHGYINGGFTTPFYKOLIGKPIQLSDSEVDPELHSLVITLNDITPVLDHPCV 523
 DB 512 LAIFHRFVDAFVSVFYKMLQCKVTLQDMESMDAEYRSLVILNDITGVLDLFSV 571
 QY 524 ENHAFGRILQHEIKPGRVVPVTEENKKEVRLVYVMFPMGIAQFATLOKGFNEILPQ 583
 DB 572 EDNCFSEVTTIDKRGRIEVTENKEVYDLYVM-IQRIEGRFAHFGSSELLPQ 630
 QY 584 HLKPPDQKELELIIGDLKIDLNDKSNTRLKRCVADSNIVRMFQAVETFEDERRARL 643
 DB 631 ELINVDERELESLIGLSSEIMEDMKCHKYRSYSENDQIKFMELMDWSNEKSRLL 690
 QY 644 LQFTGSTVPLQGFALQSGTAGAPRLFTTHLDANTDLPRATGFRNIDIPPEY 703
 DB 691 LQFTTISRIPVNGFKDLQSGD--GPRKFTIERA-GEPRKLPRAHTCFNRDLPPYTSK 746
 QY 704 EKLXKLTAVEETCGFAVE 723
 DB 747 KLDHKLSTAVEETIGFGE 766
 RESULT 11
 ID ABB73459 standard; Protein; 832 AA.
 XX ABB73459;
 AC ABB73459;
 XX 30-JAN-2003 (first entry)
 XX DE Candida albicans essential protein SEQ ID NO 7296.
 XX KM Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
 XX KM signal transduction; DNA replication; cell division; growth;
 XX KM proliferation; Candida albicans; fungicide; antifungal.
 OS Candida albicans.
 PN WO200253728-A2.
 PD 11-JUL-2002.
 PF 26-DEC-2001; 2001MO-US49486.
 PR 29-DEC-2000; 2000US-259128P.
 PR 20-FEB-2001; 2001US-0792024.
 PR 22-AUG-2001; 2001US-314050P.
 PA (ELIT-) ELITRA PHARM INC.
 PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KU;
 XX WPI; 2002-566694/60.
 DR N-PSDB; ABZ3209.
 PT Constructing strains for identifying gene products as effective targets
 PT for therapeutic intervention, by inactivating in the strain one allele
 PT of a gene and placing other allele of the gene under conditional
 PT expression -
 PS Claim 44; SEQ ID NO 7296; 167pp + Sequence Listing; English.
 XX The invention relates to constructing (M1) a strain of diploid fungal

CC cells in which both alleles of a gene are modified, comprising modifying
 CC one allele by insertion or replacement by a cassette having an
 CC expressible selectable marker and modifying other allele by
 CC recombination, of a promoter replacement fragment with a heterologous
 CC promoter, so that expression of the second allele is regulated by the
 CC promoter. (M1) is useful for constructing a strain of diploid fungal
 CC cells in which both alleles of a gene are modified. The diploid fungal
 CC cells having both alleles modified are useful for identifying a gene that
 CC is essential to the survival or growth of a fungus, a gene that
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene
 CC that contributes to the resistance of a diploid fungus to an antifungal
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus
 CC and for identifying a therapeutic agent for treatment of a mammalian
 CC disease. (M1) is useful for identifying a compound which modulates the
 CC activity of a gene product, preferably enzymatic activity, carbon
 CC compound catabolism, biosynthetic, transporter, transcriptional,
 CC translational, signal transduction, DNA replication and cell division
 CC activity. The method is useful for identifying a compound having the
 CC ability to inhibit growth or proliferation of *C. albicans* cells and for
 CC treating infection by *C. albicans*. The present sequence is that of an
 CC essential *Candida albicans* protein used in the method of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office.

CC Sequence 832 AA;

Query Match 37.0%; Score 1436.5; DB 23; Length 832;

Best Local Similarity 36.4%; Pred. No. 1,36-111; Indels 167; Gaps 21;
 Matches 326; Conservative 124; Mismatches 233;

QY 7 RLTVLCAKNIKKKQAGFLGCVRL-LSVAISRLMDTGVOGL-DLCLNLSDDDAIR 66
 DB 17 INKVALESILYKRDVFRQDPFAVLTVYDGS-QTKITITAKKTINPYMNETFN-FOAKED 74
 QY 67 SI-TISVNNHKKIKHKKQAGFLGCVRL-LSVAISRLMDTGVOGL-DLCLNLSDDDAIR 122
 DB 75 SILVQVFDQKKF-KKQDQFLGVINRIGVIDLSLNSSEBITTRDLKSN--EVLAAVS 131
 QY 123 GQIVVSLQ-----TRD-----RLGTGSAVVDCKGLLE 149
 DB 132 GKIIIVISHNNSNGGVTTATTTGTGASSNNIATITSGVNNLRISATTTANSTAAQS 191
 QY 150 NEGTV-YEDSGPCR---PLSCMEEPAYTSTGAAGGNCRCVZESPDDQLQQRRL 205
 DB 192 SDATVANGANGSGPTSLPPIGQGHPEPTATPGAGAAAGASROYSPEDQYGLPPGWER 251
 QY 206 NPDVAG-----SLQTPNCRPHGOSP----- 226
 DB 252 RTDNFGRITYVDHNRRTTQGRPALHOSERTERGQROSEFBAERQRKRTLGEGSVSP 311
 QY 227 -----ELPEGYEQRTVQGOY 242
 DB 312 LPTGSGNSITSGNTVNASGANTPVNPAAYVMAASGATTSGLDELPSGEGEORFTTEGRP 371
 QY 243 YELHTQTGVTWHDRIPIRDLNSVY-----CDELGPLPQMEYRSTVSGIYVCH 293
 DB 372 YFVDHNRITTTWVDRRQOYIRTEGPNTTIOQYPSQLGPIPSMEKRLNTRARVFDH 421
 QY 294 NNRTTQFTDPRLLHIMNHQCOLKEPSQPLPSPSGSLDEELPAQRYERDLVQKLKYLH 353
 DB 432 NRTKTTWDDPR-----LPS--SL-DQVNP--QYKRPFRFRVYFRS 467
 QY 354 ELSLQPGQAGHCRIVSREEIFEESYQIMKAPVDLKKRLMVFREBEGLDGQVAREW 433
 DB 468 QPAL-RILPGCHIKVRDHI FEDSYQIMKOTPEDLKKRLMIFDGBEGLDGQVAREW 536
 QY 414 LYLICHEMLNPFYGLFOYSTDNIVMLQINPDS-INPDLSYFHFVGRIMGLAVHGYIN 473
 DB 527 PFLSHDMENPFYCLFEVSSHNDNYTLQINPSGINPEHLNFKFGRVGLGVHRRFLD 586
 QY 474 GGFYVFFYKQLGKPIQLSDLESVDPRLKSLVWLTENDIPVLDHFFCVENHAFGRILQ 533

DB 587 APFVGLYKMLKHKVIVLQDMGVDAEFRLSKNLTINDITGILDITSAEESGZIVE 646
 QY 534 HEKENGNAVPTTEENKKEYVRLVYWRMREIGIEAQLAQGFNELLPOHLKFPDOKE 593
 DB 647 VDLKPGARDIEVEENKHEVELEITMRISKEVEQFKAIFGFNELLPOELVNFDDRE 706
 QY 594 LELIGGLDIDLDIMDKSNTRLGHCVAASNIVAMPQAVETFEERRRRLILOFTVSTRV 653
 DB 707 LELIGGLAIEDDMKKHIDYRGVENQVIAQWTKCNENDSOKARLLOFTGTGRI 766
 QY 654 PLOGFKALQSGTGAAPRLFTIHLDPANTDNPXAKTCFNRIDIPYSEYKELKTLTA 713
 DB 767 PVANGFKDLQSSD---GPRFTIEKA-GEANQLPKSHTCFNRVLDPLPYDYESTKXLTILA 822
 QY 714 VETTCGPAVE 723
 DB 823 VETTVGFGQE 832

RESULT 12
 ABJ26104
 ID ABJ26104 standard; Protein; 869 AA.
 AC ABJ26104;
 DT 16-APR-2003 (first entry)
 DE Aspergillus fumigatus essential gene protein #762.
 XX Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
 XX cancer; contamination; biofilm; antibody; immune response.
 XX Aspergillus fumigatus.
 XX WO200286090-A2.
 XX 31-OCT-2002.
 XX 23-APR-2002; 2002WC-US13142.
 XX 23-APR-2001; 2001US-285697P.
 XX 27-APR-2001; 2001US-287066P.
 XX 05-JUN-2001; 2001US-295890P.
 XX 09-JUL-2001; 2001US-303899P.
 XX 31-AUG-2001; 2001US-316362P.
 XX (ELIT-1) ELITRA PHARM INC.
 PI Jiang B, Tiehoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
 DR WPI; 2003-093124/08.
 XX New purified or isolated nucleic acids of essential genes of
 PT Aspergillus fumigatus, useful for treating or preventing infections by
 PT A. fumigatus, or for treating a non-infectious disease in a subject
 PT e.g. cancer -
 XX
 PS Disclosure; Page -: 175pp; English.

CC The invention relates to novel purified or isolated nucleic acids of
 CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
 CC the invention are used to treat or prevent infections by a pathogenic
 CC organism such as *A. fumigatus*, to treat a non-infectious disease in a
 CC subject (e.g. cancer), to prevent or contain contamination of an object
 CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
 CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for
 CC expressing recombinant protein for characterization, screening or
 CC therapeutic use, as markers for host tissues in which the pathogenic
 CC organisms invade or reside, for comparing with the DNA sequence of *A.*
 CC *fumigatus* to identify duplicated genes or paralogues having the same or
 CC similar biochemical activity and/or function, for comparing with DNA
 CC sequences of other related or distant pathogenic organisms to identify
 CC potential orthologous essential or virulence genes, for selecting and

CC invention.
 XX Sequence 854 AA;
 SQ Query Match 33.9%; Score 1318; DB 20; Length 854;
 Best Local Similarity 35.4%; Pred. No. 6.8e-120;
 Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;

QY 6 KIRLVLCAR-NLAKDFFLLPDPFAKIVDGGCHSDTYKNTLDPMMCHDLYGK 64
 DB 9 QLOITVISATLKNKNKWF-G-PSPIVEVVD--GOSKTEKCNNTNSPKMKOPLTVITP 65
 QY 65 TDSITISVNNKKIKKKOGAGFLGCYRL-----LSNAISRLKDTGYCR----- 107
 DB 66 TSKLCFRVMSHQLT---KSDVLLGTAGLDIYETLKSNMKLEBVTMTQLVADKKEPTEFM 122
 QY 108 --LDLCKLNSPDIDANVGQIVS-----LQTD--RIGTGS--VDCR 145
 DB 123 GDLSTVC---LDGLQVFAEVVINGETSCESTTQNDGCRITDDTRVSTNGSEDEPVAS 178
 QY 146 G-----LLENQTVYEDSGGRPLASCFMEBPAYT-----DSTG 179
 DB 179 GENGRANGNNSPISLNG--FKSRRPRP-----SRPPPTRRPASVNGSPSTSDSDG 231
 QY 180 AA-----AGGNCRFVESPSQ-----DQRLQAR 203
 DB 232 SSTSLPPTNTVNTSTSEGTSLIIPLTISGSGSPRLNTVSQLPPLPGWEQRY--- 287
 QY 204 LRNPVAGSL-----QTPONRPHGHSPLPEGYBORTVQGOVYFLHTQGVSTWH 255
 DB 288 ---DQGRVYVYDVAEKITWDPE---PLPGWEKRVDMKRIYVDHFTRTTWQ 338
 QY 256 DPRIP-----BDL-----NSVNCDELGPLPGWEVR 281
 DB 339 RPTLESVRNTEQWOLQSGQAMQCFQRFYIGNODLPATSONKEFDLGPLPGWEKR 398
 QY 282 STVGRIYFDHNNRTQFTDPRLIHNMNQCCKSPQPLP----- 324
 DB 399 TDSNGRVYFNHNRITQWEDPR-----SQGLNE--KPLPGEMERTVDGITYFVDH 450
 QY 325 -----PSEGLEDEELPAQRYERDLYQKLVLR-----HEVSLQCPAGHCRIEVSRE 372
 DB 451 NRRATYIDPRTGSAIDNGPQIAYVDFKAKYQYFRFMOQLAMPQ---HIKITVRK 506
 QY 373 RIFESYRQIMKMRPKLKKRLMVKPRGEGLDYGVAREMYLLCHEMLNYYGLFOYS 432
 DB 507 TLFEDSFQOIMSFQDLRRRLMWIFPGEGLDYGVAREWFLLSHEVLNPMYCLFEYA 566
 QY 433 TDNIYMLQINPDSSINDHLSEFHFVGRIMGLAVFHGYINGFTVPFYKQDLGKPIQS 492
 DB 567 GKONYCQINPNASTINDHLKTPFRIGRTAMALFPGKFLIDGFSLPFYKRIILNKPVLK 626
 QY 493 DLESVDPELHKSILWILENDITPV-LDHTCVENHAFGLQHELKPNRGNPVTEBKK 551
 DB 627 DLESIDPEFYNLSLWVENNIEECGEMYSVXKEILGKSHDLKPNQGNLVEENKE 686
 QY 552 EYVRLYVMMRFMRGIEQFLALQGFNELIPOHLKPFDOKEHLIIGGLDKIDNDMS 611
 DB 687 EYRMVAMERLSRVEGQTOAFEGFRIILPOQYLOVFAKKEIVLLCMQETIDANDMR 746
 QY 612 NTRLKGVADSNIVRFWQAVETFDERRARLLQPVTSGRVPLQGFKALQSGTGAAGR 671
 DB 747 HAIRHYRTSKQIMFWQFVKEIDNEKEMRLIQPVGTCTELVGGFADIMSN--GPO 803
 QY 672 LFTIHLIDANTDNPKAHTCFNRIDIPYSYKLYEKLITAVEECGPAVE 723
 DB 804 KFCIEKV-GKENMLPRSHTCNRLDLPYASYEQLEKLLFAIEETGEGQ 854

AC AAY0948;
 XX 21-OCT-1999 (first entry)
 DE Human E3 ubiquitin protein ligase protein.
 XX E3 ubiquitin protein ligase; h-E3 UPL; antiinflammatory; antidiabetic;
 KW immunosuppressive; neuroprotective; cytoskeletal; antiarthritic; cardiant;
 KW immunomodulator; antiviral; treatment; screening; gene therapy; cancer;
 KW inflammation; autoimmune disease; neurological disease; apoptosis;
 KW endothelial cell; proliferation; differentiation; angiogenesis; cachexia;
 KW peripheral vascular disease; hemopoietic disorder; arthritis; leukemia;
 KW pulmonary disorder; diabetes; viral infection; human.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Domain 275..306
 FT /label= MWI
 FT /note= "WW protein interaction domain I"
 FT Domain 307..340
 FT /label= MWII
 FT /note= "WW protein interaction domain II"
 FT Domain 386..420
 FT /label= MWIII
 FT /note= "WW protein interaction domain III"
 FT Domain 427..460
 FT /label= MWIV
 FT /note= "WW protein interaction domain IV"
 FT
 FN NO9940201-AI.
 XX
 PD 12-AUG-1999.
 XX
 PF 02-FEB-1999; 99WO-GB00353.
 XX
 PR 30-APR-1998; 98US-0070060.
 PR 05-FEB-1998; 98US-0073839.
 XX
 PA (ZENB) ZENBICA LTD.
 XX
 PI Ghildyal N, Husted CM;
 PI WPI: 1999-508506/42.
 DR N-PSDB; AA09235, AAY09236.
 DR
 XX New human proteolytic accessory enzyme and its modulators useful
 PT for treating disease conditions like inflammation or autoimmune
 PT diseases
 PS Claim 1; Fig 3; 95pp; English.
 XX
 CC This invention describes a novel human polynucleotide (I) which encodes a
 CC E3 ubiquitin protein ligase, h-E3 UPL (II). The products of the invention
 CC have antiinflammatory, immunosuppressive, neuroprotective, cytoskeletal,
 CC antiarthritic, immunomodulator, antidiabetic, antiviral and cardiant
 CC activity. The products of the invention can also be used for treating
 CC patients with disorders mediated by the biological and/or pharmacological
 CC activity of h-E3 UPL. The nucleic acid sequences encoding E3 UPL are used
 CC in expression systems as assay for agonists and antagonists for the E3
 CC UPL protein. The E3 UPL protein is used in screening assays to identify
 CC blockers and antagonists. They are also used in gene therapy. Specific
 CC modulation of biological and/or pharmacological activity of novel h-E3
 CC UPL via administration of a modulator or heterologous expression, is used
 CC for treating physiological conditions like inflammation, autoimmune
 CC diseases, neurological disease, apoptosis, endothelial cell physiology
 CC (e.g., proliferation, differentiation), peripheral vascular disease,
 CC angiogenesis, cancer, hemopoietic disorders, arthritis, cachexia,
 CC leukemia, pulmonary disorders, diabetes and viral infection. The
 CC nucleotide sequences which encode h-E3 UPL may also be employed in
 CC analysis to map chromosomal location e.g., screening for functional
 CC association with disease markers. They are also used as screening tools
 CC in the identification of appropriate human subjects and patients for

therapeutic clinical trials. The sequences can also be used to detect the presence of the mRNA transcripts in a patient or to monitor the modulation of transcripts during treatment. This sequence represents the human E3 ubiquitin protein ligase protein described in the method of the invention.

Sequence 852 AA:

Query Match 33.5%; Score 1301; DB 20; Length 852;
Best Local Similarity 34.7%; Pred. No. 3.2e-118;
Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

6 KRLVLCAL-NAKKDFRLDPRAKIVVDGSGCHSTDTVKTTLDPKNQHYDLVYK 64
9 QLOITVISAALKENKKNKMGF-PSPYEVAVD--GOSKTEKKNMNSPKMQLVITVP 65
65 TDSITISVNNHKKHK--QGAGPFCVRLSNASRLKDT-----GYGK-----L 108
66 VSKLHRVMSHQTLSKDVLLGTAALDIYETLSNNMKLEVVVTLQLGSKPEPTTIGDL 125
109 DLCKLNPSTDAVRGQIVLSLT-----HDIRIGTGS-----VYDC 144
126 SIC---LDGLQLESVTVNGETTESASQNDGSRKMDETRVSTNGSDPDAGAGEN 181
145 RGL-----LENEGTVYEDSGRPPLSCFMEEPAPYDTSGAAGGNCFFVS----- 192
182 RRVSGNNSPGLSNG--FKPSRPFP--SRPPPTPRRPASVNGSPSATSSDSST 234
193 -----PSCDQRLQAQRLNPD 208
235 GSLPPTNTNTSEAGTGLIPLTISGSGSPRLNPVTAQLPFGMQRV-----D 286
209 VRSSL-----QTQNRPHGHQSPLEPEGYEORTVQGVYFHTGTGSTVHDPRI 260
287 QGGRVYVYDVHEKRTWDRPE--PLPPGWERVDNNGRIYYVDHRTKRTTQRPFLB 341
261 -----RDL-----NSVNCDELGPFPPEMEVSTVSG 286
342 SVANVEGMOQRSQLGAMQGFQRIYGNQDLFATSQSKEDPLGRLPPMEKRTDSNG 401
287 RIFVVDHNNFTTFTDPRLLHMHHCQLKEPSQPLP----- 324
402 RYFVNNHTRTITWEDR-----SQQLNE--KPLPEGEMKRTVDSIPYFVDHNRRT 453
325 -----PSEGLSEDEELPAORYERDLVOKLVLR--HELSLOPQAGHRIEVSREIPEE 377
454 TYIDPRGKALDNGQIAYVDFRAKQVYFRKQQLAMPQ---HKTIVTKRTLPED 509
378 SYRQIMKRRPKDKKLMVYKFRGEGSLDYGVAREMLYLCHENLNPYYGLPQSTDNLY 437
510 SFOQIMSFSPQDLRRRLWVLPFGEGLDYGVAREVFFLSHEVINPMVCLFEYAGKDN 569
438 MLOINPDSSINPDHLSYFHFVGRIMGLAVFHGNYNGSTVFYKQLIGKPLQSDSESV 497
570 CLOINPASVYNPDHLLKTFRTIGFIMALPHGKFTIDTGSLSFFYKRLANKPYGLKDL 629
498 DPELHSLVWILNDITPV-LDHTFCVENNAFRLLQHELEKNGNVPYTEENKKEYRL 556
630 DPERYNLIWKENNIEECLEWYFSDVKEILGEIKSHDLKNGNLIWTEENKEEYIRM 689
557 VAMRFRGIEAGFALQKGFNELLPOHLKPPDOKEILLIGLDKIDIDMKSNTRLK 616
690 VAEWRUSRGVEBQTQAFBGFENILPQOYLQYFDALKELEVLGQHOEDLNDWQHAIYR 749
617 HCVADSNIVFMQAVETPDEERRARLLQFVVGSTFVPLQGFKALQGSTGAGPRLFTIH 676
750 HYARTSKQIMFMQVAYEIDNEKRMKLLQFVVGTRCLPVGFGADLMGNS--GPOKFCIE 806
677 LIDANTDNLPKATCPRRIDIPYSEYELVETLTLAVETGFAVE 723
807 KV-GKENWLPFRSHTCFRDLDPYKSYEQLKEXLLFALETEGFOE 852

RESULT 15
AA25170
ID AA25170 standard; Protein: 927 AA.
XX
AC AA25170;
XX
DT 09-SEP-1999 (first entry)
XX
DE Human KIAAN ligase protein fragment.
XX
KW Inhibitor; ubiquitin-mediated proteolysis; IkappaB polypeptide; HECT;
XX ubiquitin-conjugating system; homologous to E6-AP carboxyl terminus;
XX ligase; ubiquitin; ubiquitination; RSC ligase; KIAAN ligase; UBC4;
XX MAD3; modulator; treatment; proliferative disorder; apoptosis; sepsis;
XX differential; disorder; viral infection; tissue wasting disorder;
XX cachexia; malignancy; inflammatory disease; parasitic disease;
XX tuberculosis; IL-2 therapy; rheumatoid arthritis; rheumatoid spondylitis;
XX osteoarthritis; gouty arthritis; respiratory distress syndrome;
KW cerebral malaria; chronic pulmonary inflammatory disease; silicosis;
KW pulmonary sarcoidosis; bone resorption disease; reperfusion injury;
KW graft versus host reaction; allograft rejection; Crohn's disease;
KW ulcerative colitis; pyresis; multiple sclerosis; autoimmune diabetes;
KW systemic lupus erythematosus; leprosy; AIDS.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..927
FT /note="partial sequence, no start codon given"
XX
XX MO9904033-A1.
PD 28-JAN-1999.
XX
PF 16-JUL-1999; 98MO-US14638.
XX
PR 16-JUL-1997; 97US-0895601.
XX
PA (MTO-) MINOTIX INC.
PI Beer-Romero P, Glaes SJ, Rolfe M, Strack PR;
PI WPI; 1999-132274/11.
DR N-PSDB; AA278494.
XX
PT Identifying modulators of I-kappa-B proteolysis - used to develop
PT products for treating e.g. proliferative and/or differentiative
PT disorders, infections, tissue wasting, cachexia or AIDS
XX
PS Disclosure; Page 68-71; 79pp; English.
XX
CC This invention describes a novel assay identifying an inhibitor of
CC ubiquitin-mediated proteolysis of an IkappaB polypeptide. The method
CC comprises: (a) providing a ubiquitin-conjugating system including the
CC IkappaB polypeptide and a HECT (homologous to E6-AP Carboxyl Terminus)
CC ligase and ubiquitin, to promote ubiquitination of the IkappaB
CC polypeptide by the HECT ligase; (b) contacting the ubiquitin-conjugating
CC system with a candidate agent; (c) measuring a level of ubiquitination of
CC the polypeptide in the presence of the candidate agent and (d) comparing
CC the measured level of ubiquitination in the presence of the candidate
CC agent with ubiquitination of the IkappaB polypeptide in the absence of
CC the candidate agent; where a statistically significant decrease in
CC ubiquitination of the IkappaB polypeptide in the presence of the
CC candidate agent is indicative of an inhibitor of ubiquitination of the
CC IkappaB polypeptide. The assays can be used to identify compounds which
CC modulate binding and/or ubiquitination of an IkappaB (or other cellular
CC or viral substrate) by a HECT ligase, such as RSC or KIAAN. Such
CC modulators can be used e.g. in the treatment of proliferative and/or
CC differentiative disorders, to modulate apoptosis in the treatment of
CC viral infections, and in the treatment of tissue wasting disorders e.g.
CC cachexia secondary to infection or malignancy, cachexia secondary to
CC human AIDS, inflammatory diseases, parasitic diseases, tuberculosis and
CC high dose IL-2 therapy; rheumatoid arthritis, rheumatoid spondylitis,

CC osteoarthritis, gouty arthritis and other arthritic conditions, sepsis,
CC respiratory distress syndrome, cerebral malaria, chronic pulmonary
CC inflammatory disease, silirosis, pulmonary sarcoidosis, bone resorption
CC diseases, reperfusion injury, graft versus host reaction, allograft
CC rejections, Crohn's disease, ulcerative colitis, or pyrosis, in addition
CC to a number of autoimmune diseases such as multiple sclerosis, autoimmune
CC diabetes, systemic lupus erythematosus; and ENL in leprosy, HIV, and
CC AIDS. This sequence represents a human KIAAN ligase which is used in the
CC method of the invention.

XX Sequence 927 AA;

Query Match 32.4%; Score 1259; DB 20; Length 927;
Best Local Similarity 33.6%; Pred. No. 5.1e-114;
Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;

QY 7 IRLTYCAKVLAKKDPFRLLDPFAKIV---VDGGQCHSTDTVANTLDPKKN----- 55
DB 48 VRVAVIAGIGLAKKDLIGASDPYVTVLYPDMGVLTISVQTKIKSLNPKMNEILFRV 107
QY 56 ---QH-----YDLVYKTD--SITISVWN-----HKIKKKOG 83
DB 108 HPQOHRLLPEVPEBNELTRDDPLGQVDVPLPLPTENRLEKPYTFKDPVLAHPSHKSRY 167
QY 84 AGFLGCVRLISNAISRLKDT-----GYQLD---LCKL----- 113
DB 168 KGYL-----RLKKTLYLPKTSGEDNAEQAELEPEGMVVLDPDACHLQOQOE 216
QY 114 -----NPSD--TDVARGIVSLQ---TRDR 134
DB 217 PSPLPPEWEERODILRRTTYVNHESRRTQWKRPPTPODNLTDANGNIQOQARAFTTRQ 276
QY 135 IGTGGSVDCRGLEN-----EGTVYEDSG-PGRPLSCFMEBPAYTD----- 176
DB 277 ISEETESVDNQESSEWEIIRDEATMYSSQAFSPSPSSNLDVPTHLAEINARLTIFG 336
QY 177 ----STGAAAGGNCR-----FVESPS-----QDQRLQARL-- 204
DB 337 NSAVSQPASSNSHSSRGSLOAYTFEEQPTLPVLLPTSSGLDPGWEKQDERGRSYVDH 396
QY 205 -----RNPVNGSLQTPQ-----NRPHGQSPLEPEGYEQ 234
DB 397 NSRTTWTKPTVATVETISQLTSSQSSAGPQSQASTSDGQGVTPQPSLEQGFPEKWEV 456
QY 235 RTTVQGVVFLHTQGVSTHDP--IPRDLASV---NCDELGPLPGEWEYRSTVSGRI 288
DB 457 RHAENGRPFIDNITKTWEDPRKIPALHKGKTSLDTSNDLGPLPGEWEERTHTDRI 516
QY 289 YFVDHNNRTQFTDPRLHIMNHQCOLKEPSOPLPLPSEGSLEDEELPAQRYERDLVOKL 348
DB 517 FYINHNKIKTQWEDPRLENV---AITGPAYP-----YSRDYKRY 553
QY 349 KVLRLHEISLQOPAGHCRLEVSREELFEESYRQIMKRPXD-LKRLMVKFRGEGLDYG 407
DB 554 EFERKRLKKQNDIPNKEFMKLRRATVLEDYSYRIMGVKRADELKARLWIEFDEKGLDYG 613
QY 408 GVAHEMLYLCHMLNPPYGLFOYSTDNLYMLOINPDSSI-NPDHLYFHFVGRIMGLAV 466
DB 614 GVAHEMFLLSKENENFYGLFEYSATDNTYLQINPSGICNEHLSYFKFGRVAGAV 673
QY 467 FGHYINGGFTVPFYKQLGKPIQLSDLSVDEBELKSLWILENDITPVLDTEFCVEHN 526
DB 674 YHGILDDGFIRPFYKWLKPKITLHDMESVDSEYNSLRWILENDPTE-LDLRFIIDE 732
QY 527 AFGRILOHELKPRGRNVPVTEENKKEYVLLYVMPRMGIEAOFLAQGFNELIPOHL 586
DB 733 LFGQTHQHELKNGSEIVTNKKNKEIYLVIQMRFVNRIQKMAAFKGGFELLPODLI 792
QY 587 KPEDQKELEHIIIGLDKIDLNDKSNTRLNG-CVADSNIVRFWQAVETFEDEERARLLQ 645
DB 793 KIFDENLELELMGLGVDVNDVREHTKYNGYSANHOVIQWFVKAVLMMOSEKRIRLLQ 852
QY 646 FVTGSTRVPLQGFRLAGSTGAAGPRLFTIHLIDANTDLPKAHTCFNRIIDIPYESTEK 705

DB 853 FYTGSRVPMNGFAELIYSN---GPGSFIVGQW-GTEPKLPRAHTCFNRDLDPYSSFEE 908
QY 706 LYEKLLTAVESTCGF 720
DB 909 LMDKLQMAIENTQGF 923

Search completed: February 20, 2004, 15:28:29
Job time : 44.3032 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 15:26:11 ; Search time 15.2366 Seconds

(without alignments)
2007.718 Million cell updates/sec

Title: US-10-009-945-2

Perfect score: 3884

Sequence: 1 GGSIRKRLTVCAKYLAKK.....ELTYEKLTAVETGFAVE 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCtUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2913.5	75.0	735	3	US-08-539-205A-2
2	2913.5	75.0	735	4	US-08-392-163A-2
3	1471	37.9	766	3	US-08-539-205A-4
4	1471	37.9	766	4	US-08-392-163A-4
5	1318	33.9	854	2	US-08-070-060-4
6	1318	33.9	854	2	US-08-357-746-4
7	1301	33.5	852	2	US-08-070-060-3
8	1301	33.5	852	2	US-08-357-746-3
9	1259	32.4	927	3	US-08-895-601-6
10	1228.5	31.6	834	3	US-08-539-205A-6
11	1228.5	31.6	834	4	US-08-392-163A-6
12	1225	31.5	906	3	US-08-630-916A-48
13	1047	27.0	683	3	US-08-630-916A-46
14	530.5	13.7	874	2	US-08-247-904B-8
15	530.5	13.7	874	2	US-08-767-942A-21
16	509	13.1	866	2	US-08-100-692-1
17	509	13.1	866	2	US-08-674-030-1
18	473	12.2	1083	3	US-08-885-601-5
19	241	6.2	486	3	US-08-348-518C-2
20	238.5	6.1	472	3	US-08-348-518C-5
21	238.5	6.1	472	3	US-08-476-509B-5
22	162.5	4.2	54	3	US-08-630-916A-124
23	162	4.2	448	3	US-08-476-509B-2
24	161.5	4.2	454	3	US-08-348-518C-4
25	154.5	4.0	454	3	US-08-476-509B-4
26	154.5	4.0	55	3	US-08-630-916A-75
27	148.5	3.8	58	3	US-08-630-916A-84

28	148	3.8	51	3	US-08-630-916A-117	Sequence 117, App
29	147.5	3.7	54	3	US-08-630-916A-74	Sequence 74, App
30	145	3.6	51	3	US-08-630-916A-73	Sequence 73, App
31	141	3.6	54	3	US-08-630-916A-36	Sequence 36, App
32	140	3.6	54	3	US-08-630-916A-118	Sequence 118, App
33	139	3.6	51	3	US-08-630-916A-116	Sequence 116, App
34	138	3.6	1105	3	US-08-710-249-2	Sequence 2, App
35	138	3.6	1105	4	US-08-220-157A-2	Sequence 2, App
36	133	3.4	224	3	US-08-630-916A-50	Sequence 50, App
37	132	3.4	51	3	US-08-630-916A-115	Sequence 115, App
38	130	3.3	38	3	US-08-630-916A-24	Sequence 24, App
39	130	3.3	38	3	US-08-630-916A-26	Sequence 26, App
40	130	3.3	38	3	US-08-348-518C-15	Sequence 15, App
41	130	3.3	38	3	US-08-348-518C-18	Sequence 18, App
42	130	3.3	38	3	US-08-476-509B-15	Sequence 15, App
43	130	3.3	38	3	US-08-476-509B-18	Sequence 18, App
44	130	3.3	335	2	US-08-844-312-2	Sequence 2, App
45	128	3.3	38	3	US-08-630-916A-25	Sequence 25, App

ALIGNMENTS

```
RESULT 1
US-08-539-205A-2
Sequence 2, Application US/08539205A
Patent No. 6001619
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Nelsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,205A
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-539-205A-2
Query Match 75.0%; Score 2913.5; DB 3; Length 735;
Best Local Similarity 72.9%; Pred. No. 1.4e-280;
Matches 552; Conservative 63; Mismatches 71; Gaps 9;
```

5 IKRLTVCAKYLAKKDPFRLLDPFAKIVVDSGGCHSTIVKNTLDRKMQHYDLYGK 64
12 VKRLT-----GLPDPFAKVVVDSGGCHSTIVKNTLDRKMQHYDLYGK 58
65 TDSITTSVNMHKIKKQAGFGCVRLISNLSLKXGVRLLCLNPSDPAVVGQ 124

Db 59 SDSVTISVWNNKKIKHKKQAGFLGCVRLLSNAINRLKDTGYORLDCLGPNNDTVRGQ 118
 QY 125 IVVSLQTRDRIGTGSSVYDCRGLLENE-----GTVY----- 155
 Db 119 IVVSLQSRDRIGTGSSVYDCRGLLENE-----GTVY----- 178
 QY 156 --EDSGRPLSCFMEBPAYTSTGAAGGNCRFVESPQDQRLQARLNPDVGS 213
 Db 179 ASFYSSPGRPLSCFVDEMTPISTNATCG-----QSDPRLARRRARSQHRMYM 229
 QY 214 QTPQNRPHQSELPBEGYEQRTTVQGVYFLHTQGVSTWHPRIIPDLNSVNCDEIGP 273
 Db 230 ----SRTHHTPPDLPEGYEQRTTVQGVYFLHTQGVSTWHPRIIPDLNSVNCDEIGP 285
 QY 274 LPPGWEVRSVSGRIYFVDHNNRTTQFTDPR----LHHIMHQCQLKEPSQPLPSEGS 329
 Db 286 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQOLKQOQOQVY----S 341
 QY 330 L--EDELPAQRYERDLYQKLVLRHLSLQOPQAGHCRIEVSREEIFEESYQIMQR 386
 Db 342 LCPDTECLTPRYKRDLYQKLVLRHLSLQOPQAGHCRIEVSREEIFEESYQIMQR 401
 QY 387 PKDLKRLMKFRGEGLDYGVAREMYLCSHEMLNPPYGLFOYSRDIYTLQINPDSA 446
 Db 402 PKDLKRLMKFRGEGLDYGVAREMYLCSHEMLNPPYGLFOYSRDIYTLQINPDSA 461
 QY 447 INPDLSYFHFVGRINGLAVFHGHYINGFTVBPYKQLGKPIQSLDESVDPELHSLV 506
 Db 462 VNPBLSYFHFVGRINGLAVFHGHYINGFTVBPYKQLGKPIQSLDESVDPELHSLV 521
 QY 507 WILNDITPVLDHFCVEHNAFGRILQHEIKNGRVNPTTEENKKEYVLYNNMFMGI 566
 Db 522 WILNDITGVLDHFCVEHNAFGRILQHEIKNGRVNPTTEENKKEYVLYNNMFMGI 581
 QY 567 EAQFLALQKGFNEILPOHLKRPDOKELELIGLKDIDLMKSNTRLKHCVASDNIVR 626
 Db 582 EAQFLALQKGFNEILPOHLKRPDOKELELIGLKDIDLMKSNTRLKHCVASDNIVR 641
 QY 627 WFOAVETFEDEERARLLQFTVSTVPLQGFALQSGTGAAPRLFTIHLIDANTDLP 686
 Db 642 WFOAVETFEDEERARLLQFTVSTVPLQGFALQSGTGAAPRLFTIHLIDANTDLP 698
 QY 687 KAHTCFNRIDIPYBSYEKLYEKLTAIETCGFAVE 723
 Db 699 KAHTCFNRIDIPYBSYEKLYEKLTAIETCGFAVE 735

RESULT 2
 US-09-392-163A-2
 ; Sequence 2, Application US/09392163A
 ; Patent No. 6503742
 ; GENERAL INFORMATION:
 ; APPLICANT: Beach, David H.
 ; APPLICANT: Caligiuri, Maureen
 ; APPLICANT: Neleky, Bradley
 ; TITLE OF INVENTION: Ubiquitin ligases, and Uses Related Thereto
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/392,163A
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/539,205
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: CSV-005.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 832-1000
 TELEFAX: (617) 832-7000
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 735 amino acids
 TYPE: amino acids
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-392-163A-2

Query Match 75.0%; Score 2913.5; DB 4; Length 735;
 Best Local Similarity 72.9%; Fred. No. 1.4e-280;
 Matches 552; Conservative 63; Mismatches 71; Indels 71; Gaps 9;

5 IKRLTVLCAKMLAKKDFRLLPDPFAKIVDGSQCHSTDTVKNLTPKNNQHYDLYGK 64
 Db 12 VKRLTL-----GLPDPFAKIVDGSQCHSTDTVKNLTPKNNQHYDLYGK 56
 QY 65 TDSITTSVNNKKIKHKKQAGFLGCVRLLSNAISRLKDTGYORLDCLGPNNDTVRGQ 124
 Db 59 SDSVTISVWNNKKIKHKKQAGFLGCVRLLSNAINRLKDTGYORLDCLGPNNDTVRGQ 118
 QY 125 IVVSLQTRDRIGTGSSVYDCRGLLENE-----GTVY----- 155
 Db 119 IVVSLQSRDRIGTGSSVYDCRGLLENE-----GTVY----- 178
 QY 156 --EDSGRPLSCFMEBPAYTSTGAAGGNCRFVESPQDQRLQARLNPDVGS 213
 Db 179 ASFYSSPGRPLSCFVDEMTPISTNATCG-----QSDPRLARRRARSQHRMYM 229
 QY 214 QTPQNRPHQSELPBEGYEQRTTVQGVYFLHTQGVSTWHPRIIPDLNSVNCDEIGP 273
 Db 230 ----SRTHHTPPDLPEGYEQRTTVQGVYFLHTQGVSTWHPRIIPDLNSVNCDEIGP 285
 QY 274 LPPGWEVRSVSGRIYFVDHNNRTTQFTDPR----LHHIMHQCQLKEPSQPLPSEGS 329
 Db 286 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQOLKQOQOQVY----S 341
 QY 330 L--EDELPAQRYERDLYQKLVLRHLSLQOPQAGHCRIEVSREEIFEESYQIMQR 386
 Db 342 LCPDTECLTPRYKRDLYQKLVLRHLSLQOPQAGHCRIEVSREEIFEESYQIMQR 401
 QY 387 PKDLKRLMKFRGEGLDYGVAREMYLCSHEMLNPPYGLFOYSRDIYTLQINPDSA 446
 Db 402 PKDLKRLMKFRGEGLDYGVAREMYLCSHEMLNPPYGLFOYSRDIYTLQINPDSA 461
 QY 447 INPDLSYFHFVGRINGLAVFHGHYINGFTVBPYKQLGKPIQSLDESVDPELHSLV 506
 Db 462 VNPBLSYFHFVGRINGLAVFHGHYINGFTVBPYKQLGKPIQSLDESVDPELHSLV 521
 QY 507 WILNDITPVLDHFCVEHNAFGRILQHEIKNGRVNPTTEENKKEYVLYNNMFMGI 566
 Db 522 WILNDITGVLDHFCVEHNAFGRILQHEIKNGRVNPTTEENKKEYVLYNNMFMGI 581
 QY 567 EAQFLALQKGFNEILPOHLKRPDOKELELIGLKDIDLMKSNTRLKHCVASDNIVR 626
 Db 582 EAQFLALQKGFNEILPOHLKRPDOKELELIGLKDIDLMKSNTRLKHCVASDNIVR 641
 QY 627 WFOAVETFEDEERARLLQFTVSTVPLQGFALQSGTGAAPRLFTIHLIDANTDLP 686
 Db 642 WFOAVETFEDEERARLLQFTVSTVPLQGFALQSGTGAAPRLFTIHLIDANTDLP 698
 QY 687 KAHTCFNRIDIPYBSYEKLYEKLTAIETCGFAVE 723
 Db 699 KAHTCFNRIDIPYBSYEKLYEKLTAIETCGFAVE 735

QY 373 EIPESYRQIMKRPKDLKRLMVKFRGEGLDYGVAREMYLLCHEMLNPFYGLFOYS 432
 DB 507 TLFEDSFQOIMSPBQDLRLRLMVIFFGEGLDYGVAREMYLLCHEMLNPFYGLFOYS 566
 QY 433 TDNIYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPYKOLLGKPIOLS 492
 DB 567 GKDNVYCIQINPDASYINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPYKOLLGKPIOLS 626
 QY 493 DLESVDPELHKSILWTLNDITPV-LDHTFCVHNAPFRILOHELKPNRANVPYTEENK 551
 DB 627 DLESVDPELHKSILWTLNDITPV-LDHTFCVHNAPFRILOHELKPNRANVPYTEENK 686
 QY 552 EYRLVYNNRFRMGIEAOFALOKGFNELIPHLLKPPDQKLELIGGLDKIDINDMKS 611
 DB 687 EYRVMVEMRLSRGVEEQTOAFEGFENEILPQOYLQYFPAKLELVLLCGMOEIDINDMQR 746
 QY 612 NTRLKHCVAODSNIVRMQVAETFEDEERARLLQFTVGTSTRVPLQGFALQOSTGAAGR 671
 DB 747 HAIYRHYTRTSKQIMFWQFVKEIDNEKMRLLQFTVGTSTRVPLQGFALQOSTGAAGR 803
 QY 672 LFTIHLIDANTDNLPRKATCFNRIDIPYSEYKLYEKLITAVEETCGFAVE 723
 DB 804 KFCIEKV-GKENWLPKSHTCFNRIDIPYSEYKLYEKLITAVEETCGFAVE 854

RESULT 6
 US-09-357-746-4
 ; Sequence 4, Application US/09357746
 ; Patent No. 6087122
 ; GENERAL INFORMATION:
 ; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
 ; FILE REFERENCE: PHM.70312.N1
 ; CURRENT APPLICATION NUMBER: US/09/357,746
 ; EARLIER FILING DATE: 1999-07-21
 ; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
 ; EARLIER FILING DATE: 1998-02-05
 ; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
 ; EARLIER FILING DATE: 1998-04-30
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO: 4
 ; LENGTH: 854
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-357-746-4

Query Match 33.9%; Score 1318; DB 3; Length 854;
 Best Local Similarity 35.4%; Pred. No. 1,1e-121;
 Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;
 QY 6 KIRLVLCAG-NLAKKDFRLLDPPAKIVYDGGQCHSTDTVKNLTDPKKNQHYDIYVGR 64
 DB 9 QLOITVISAQKAKENKKNFVG-FSPFVEYVD--GQSKTKKCNNTNSPKKQPLVIVTP 65
 QY 65 TDSITISYWNHKKIKHKKAGAGFLGCVRL-----LSNAISRLKDTGYR----- 107
 DB 66 TSKLCFRVWSHQTL--KSDVLLGAGADIVETLKSNNMLKEVYVTLQAVGKEPLETM 122
 QY 108 --LDLCKLNPSTDAVRGQIVS-----LQTRD--RIGTGG--VYDGR 145
 DB 123 GDLVVC---LDGLQVEAEVNTNSETSGESTGNDGCRDTRVSTNGSDPEVAAS 178
 QY 146 G-----LLENEGTVEDSGPGRPLSCFMEBPAPY-----DSTG 179
 DB 179 GENKRGANGNSPSLNSNG--FKPSRPPR-----SRPPPTPRRPAVNGSPSTNSDSG 231
 QY 180 AA-----AGGNGRFPVSEPG-----DRLQAGR 203
 DB 232 SSTGSLPPTNTVNTSTEGATSGIILPLTISGSGRPPLNTVSGQAPLPGWEOQV----- 287
 QY 204 LNPVVRGSL-----QTFQNRPHGHQSPELPEGYEQRTTVGQVYFHTOTGVSTWH 255

DB 288 ----DQGRVYVYDVHVKRTTWDRB-----PLPGRMERVDNMGRIYYVDHFTRTTWQ 338
 QY 256 DRIIP-----RDJ-----NSNODLGLPLPGMEVR 281
 DB 339 RPTLESVRNVEQMOQORSQLOQAMQOFNRFYGNQDLPATSONKEPDLPLPBGWEKR 398
 QY 282 STVSGRIYEVVDHNNRTQFTDRLHHIMNHQOLKEPSCQPL----- 324
 DB 399 TDSNGAVYVFNHNRITQIOMEDR-----SQQLNE--KPLPEGEMRFTVDGIPYVDH 450
 QY 325 -----PSESLEDELPRQYERDPLVQKATLR--HELSIQOQACHCIVSRE 372
 DB 451 NRDATYIDERTGKSLDNGPQIAYRPFKAVQYFRWCQQLAMPQ--HikitvRK 506
 QY 373 EIPESYRQIMKRPKDLKRLMVKFRGEGLDYGVAREMYLLCHEMLNPFYGLFOYS 432
 DB 507 TLFEDSFQOIMSPBQDLRLRLMVIFFGEGLDYGVAREMYLLCHEMLNPFYGLFOYS 566
 QY 433 TDNIYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPYKOLLGKPIOLS 492
 DB 567 GKDNVYCIQINPDASYINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPYKOLLGKPIOLS 626
 QY 493 DLESVDPELHKSILWTLNDITPV-LDHTFCVHNAPFRILOHELKPNRANVPYTEENK 551
 DB 627 DLESVDPELHKSILWTLNDITPV-LDHTFCVHNAPFRILOHELKPNRANVPYTEENK 686
 QY 552 EYRLVYNNRFRMGIEAOFALOKGFNELIPHLLKPPDQKLELIGGLDKIDINDMKS 611
 DB 687 EYRVMVEMRLSRGVEEQTOAFEGFENEILPQOYLQYFPAKLELVLLCGMOEIDINDMQR 746
 QY 612 NTRLKHCVAODSNIVRMQVAETFEDEERARLLQFTVGTSTRVPLQGFALQOSTGAAGR 671
 DB 747 HAIYRHYTRTSKQIMFWQFVKEIDNEKMRLLQFTVGTSTRVPLQGFALQOSTGAAGR 803
 QY 672 LFTIHLIDANTDNLPRKATCFNRIDIPYSEYKLYEKLITAVEETCGFAVE 723
 DB 804 KFCIEKV-GKENWLPKSHTCFNRIDIPYSEYKLYEKLITAVEETCGFAVE 854

RESULT 7
 US-09-070-060-3
 ; Sequence 3, Application US/09070060
 ; Patent No. 5976849
 ; GENERAL INFORMATION:
 ; APPLICANT: Huestad, Carolyn M.
 ; TITLE OF INVENTION: Human E3 Ubiquitin Protein
 ; TITLE OF INVENTION: Ligase
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: ZENCA Pharmaceuticals, Inc.
 ; STREET: 1800 Concord Pike
 ; CITY: Wilmington
 ; STATE: DE
 ; COUNTRY: USA
 ; ZIP: 19850-5437
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/070,060
 ; FILING DATE: 30-Apr-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/073,839
 ; FILING DATE: 05-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Higgins, Patrick H
 ; REGISTRATION NUMBER: 39,709
 ; REFERENCE/DOCKET NUMBER: PHM.70312
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: 302.886.4889
 TELEFAX: 302.886.9221
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 852 amino acids
 TYPE: amino acid
 STANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 S-09-070-060-3

Query Match 33.5%; Score 1301; DB 2; Length 852;
 Best Local Similarity 34.7%; Pred. No. 5.4e-120;
 Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

Y 6 KIRLVLCAL-NLAKKDFRLDPDPKAVVDGSGQCHSTDTVNTLDPKWNQHYDLYVK 64
 b 9 QLOITVISAKLKENKNWFG-PSPYVEVTD--GQSKTEKCNNTNSPKWQPLTVITP 65
 Y 65 TDSITISVNHKKIHK--QGAGFLGCVRLLSNAISRLKDT-----GYQR-----L 108
 b 66 VSKLHFRVWSHQTLSKSDVLLGTAALDIYETLKNMKNLEVVVTLQLGDKKEPTETIGDL 125
 Y 109 DLCKLNPSDTPAVRGQIVVSLQT-----RDRIGTGGS-----VVDG 144
 b 126 SIC---LDGLQLESEVNTGETTCSASQNDGSGSKDETRVSTNGSDDPEDAGAGEN 181
 Y 145 RGL-----LENEGTVYEDSGPRPLSCMEEPAPYDSTGAAGGNCRFVES----- 192
 b 182 RRVGNNPSLSNGG--FKPSRPRP-----SRPPPTPRPASVNGSPSATESDGSST 234
 Y 193 -----PSQDORLOAQLRNP 208
 b 235 GSLPPTNTNTSEGATSGLIPLITISGSGSPRLNPVTOAPLPQWEQKV-----D 286
 Y 209 VRGSL-----QTPQNRPHGHSPELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP 260
 b 287 QHGRVYVDHVEKRTTWORPE---PLPPGWERRVDMNGRIYVDHFTTITTWQRTLE 341
 Y 261 -----RDLE---NSVNCDELGLPLPGWEVRSVSG 286
 b 342 SVNVEQWQLORSQLOGAQFNQRFYVGNQDLFATSQSKEFDPGLPLPGWEKRTDSNG 401
 Y 287 RIYFVDHNNRTTQFTDPRLLHMHNCOLKEPSQPLP----- 324
 b 402 RVYFVNHNTRITQWEDPR-----SQGLNE---KPLPEGWEMRFTVDGIPYFVDHNRRT 453
 Y 325 ----PSEGSLDEELPAQYERDLVQKLKVL---HLSLQOQOAGHCRIEVSRBIFEE 377
 b 454 TYIDPRTKSALDNGFOIAYVDFKAKYQYFRFCQQLAMPQ---HIKITVTRKTLFED 509
 Y 378 SYQIWMKRPDLKRLMVKFRGEGLDYGVAREWLYLLCHEMLNPPYGLFOYSTDNLY 437
 b 510 SFQIIMSFPQDLRRRLWVIFPGEGLDYGVAWEWFLLSHEVLNPMYCLFAYAGKDY 569
 Y 438 MLQINPDSSINPDHLSYFHFVGRINGLAVFHGYINGGFTVPFYKOLLGKPIOLSLESV 497
 b 570 CLQINPASYINDPLKYFRFGRIAMALFHGKFDITGFSPLPYKILNKPVLKDLSEI 629
 Y 498 DPBLKSLVWILENDITPV-LDHTCVENHAFGRILQHELKNGRNVPVTEENKKEYVL 556
 b 630 DPBFYNSLWVKNENIBECDELMYFSVDKEITLKGHDLKPNGNILVTEENKEYIEM 689
 Y 690 VAEWRLSRGVEEQTAFFGSGNEILLPQYQVDFAKLEVLICGQWIDLNDWQRAIYR 749
 b 617 HCVADSNVWVFNQAVETDEPRRALLQFTGTSTPLQGFALQOSTGAAPRLFTIH 676
 Y 750 HYARTSQINWFWQFVKEIDNEKRWLLIQFTGTGRLPVGFGFADLWGSN---GPOKFCIE 806
 b 677 LIDANTDNLKMAHTCFNRDLDPPEYSEKLYEKLITAVEETCGFAVE 723

Db 807 KV-GKENWLPERSHTCFNRDLDPPEYSEKLYEKLITAVEETCGFAVE 852

RESULT 8

US-09-357-746-3
 ; Sequence 3, Application US/09357746
 ; Patent No. 6087122
 ; GENERAL INFORMATION:
 ; APPLICANT: ZENECA Limited
 ; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
 ; FILE REFERENCE: PHM.70312.N1
 ; CURRENT APPLICATION NUMBER: US/09/357,746
 ; CURRENT FILING DATE: 1999-07-21
 ; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
 ; EARLIER FILING DATE: 1998-02-05
 ; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
 ; EARLIER FILING DATE: 1998-04-30
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 852
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-357-746-3

Query Match 33.5%; Score 1301; DB 3; Length 852;
 Best Local Similarity 34.7%; Pred. No. 5.4e-120;
 Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

QY 6 KIRLVLCAL-NLAKKDFRLDPDPKAVVDGSGQCHSTDTVNTLDPKWNQHYDLYVK 64
 Db 9 QLOITVISAKLKENKNWFG-PSPYVEVTD--GQSKTEKCNNTNSPKWQPLTVITP 65
 QY 65 TDSITISVNHKKIHK--QGAGFLGCVRLLSNAISRLKDT-----GYQR-----L 108
 Db 66 VSKLHFRVWSHQTLSKSDVLLGTAALDIYETLKNMKNLEVVVTLQLGDKKEPTETIGDL 125
 QY 109 DLCKLNPSDTPAVRGQIVVSLQT-----RDRIGTGGS-----VVDG 144
 Db 126 SIC---LDGLQLESEVNTGETTCSASQNDGSGSKDETRVSTNGSDDPEDAGAGEN 181
 QY 145 RGL-----LENEGTVYEDSGPRPLSCMEEPAPYDSTGAAGGNCRFVES----- 192
 Db 182 RRVGNNPSLSNGG--FKPSRPRP-----SRPPPTPRPASVNGSPSATESDGSST 234
 QY 193 -----PSQDORLOAQLRNP 208
 Db 235 GSLPPTNTNTSEGATSGLIPLITISGSGSPRLNPVTOAPLPQWEQKV-----D 286
 QY 209 VRGSL-----QTPQNRPHGHSPELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP 260
 Db 287 QHGRVYVDHVEKRTTWORPE---PLPPGWERRVDMNGRIYVDHFTTITTWQRTLE 341
 QY 261 -----RDLE---NSVNCDELGLPLPGWEVRSVSG 286
 Db 342 SVNVEQWQLORSQLOGAQFNQRFYVGNQDLFATSQSKEFDPGLPLPGWEKRTDSNG 401
 QY 287 RIYFVDHNNRTTQFTDPRLLHMHNCOLKEPSQPLP----- 324
 Db 402 RVYFVNHNTRITQWEDPR-----SQGLNE---KPLPEGWEMRFTVDGIPYFVDHNRRT 453
 QY 325 ----PSEGSLDEELPAQYERDLVQKLKVL---HLSLQOQOAGHCRIEVSRBIFEE 377
 Db 454 TYIDPRTKSALDNGFOIAYVDFKAKYQYFRFCQQLAMPQ---HIKITVTRKTLFED 509
 QY 378 SYQIWMKRPDLKRLMVKFRGEGLDYGVAREWLYLLCHEMLNPPYGLFOYSTDNLY 437
 Db 510 SFQIIMSFPQDLRRRLWVIFPGEGLDYGVAWEWFLLSHEVLNPMYCLFAYAGKDY 569
 QY 438 MLQINPDSSINPDHLSYFHFVGRINGLAVFHGYINGGFTVPFYKOLLGKPIOLSLESV 497
 Db 570 CLQINPASYINDPLKYFRFGRIAMALFHGKFDITGFSPLPYKILNKPVLKDLSEI 629

QY 498 DPELHKSJLVILENDITPV-LDHTFCVHNAPFORILOHELKPNRNPVPTVENKEYVRL 556
 DB 630 DPEFYNLSIWKNNIBECDELMVFSVDKEILGSIKSHDLKNGNLIIVTEENKEVIRM 689
 QY 557 YVWRFWRGTAEQALOKFENLIQHLKPDQKELELIIGLKDIDNDKSNTRLK 616
 DB 690 VAEWRLSRGVEEQAPFEGFNILQQYLOVFDAKELEVLCCGQEIIDLNDWQHAIYR 749
 QY 617 HCVADSNIVFWQAVTFDEERRARLLOFVTGSTRVPLQGFALQSGTGAGPRLFTIH 676
 DB 750 HYARTSQIWMFMQFVKEIDNEXKMLLQFVTGTRKLPVGGFADLMGN---GPKXCIB 805
 QY 677 LIDANTNLKANTCNRIIDIPYESYKLYKLLTAVETTCGRAVE 723
 DB 807 KV-GKNWLFASHTCFNRLDLPYKSYEQKELLLFAIEETEGFQOE 852

RESULT 9

US-08-895-601-6
 : Sequence 6, Application US/08895601
 : Patent No. 6060262

GENERAL INFORMATION:

APPLICANT: Beer-Romero, Peggy
 APPLICANT: Strack, Peter J.
 APPLICANT: Glass, Susan J.
 APPLICANT: Rolfe, Mark

TITLE OF INVENTION: REGULATION OF KAPPA B (IKB) DEGRADATION,
 TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO
 NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA

ZIP:

02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 APPLICATION NUMBER: US/08/895,601
 FILING DATE: 16-JUL-1997
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MIV-096.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 927 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

-08-895-601-6

Query Match

Best Local Similarity 32.4%; Score 1259; DB 3; Length 927;

Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;

7 IRLTVLCAKHLAKDFFRLPDPFAKIV-----VDGSGQCHSTDVTKVNLDPKWN----- 55

48 VRVRVIAIGLAKKIDLGASDPYRVTLVDPNGVLTSTVQTKIKSLNPKWNEILFRV 107

56 ---QH-----YDLVYVKTD---SITISVWN-----HKKHKQOG 83

108 HPQQRHLFEVDENRLTRDDFLGGQVDVPLPLPTENPRLERPPTFXDFVLHPRSHKSRV 167

QY 84 AGFLGCVRLLSNAISRLKDT-----GYQLD-----LCKL----- 113
 DB 168 KGYL-----RLKMTYLPKTSGEDDABQAELEFGWVLDQDPAACHLQOOQE 216
 QY 114-----NPSD-----TDVARGQIVVSLQ-----TRDR 134
 DB 217 PSLPPQWEERQDILGRYVYNHESRRITQWKPTQDNLTDANGNIQLOQACRAFTTRRQ 276
 QY 135 IGTGGSVVDCRGLLEN-----EGTVYEDSG-PGRPLSCFMEEPAPYTD----- 176
 DB 277 ISEITESVDNQESSENKEWIRREDEATMYSSQAFPPSPSSNLDVPTHLAEELNARLITFG 336
 QY 177-----STGAAAGGNCR-----FVESPS-----QDORLQARL-- 204
 DB 337 NSAVSQPASSNSHRRSGSLQAYTFEEQPTLPVLLPTSSGLPPGWEEKQDGRGYSYYVDH 396
 QY 205-----RNPDRVRSGLQTPQ-----NRPHGHQSPPELPEGYEQ 234
 DB 397 NSRTTWTWKPTVQATVETSQTSQSSAGPQSQASTSDSGQVTPQSPSEIQGFLLPKGWEV 456
 QY 235 RTTVQGVYELHTQTGVSTWHDPR--IPDLNSV-----NCDELGPLPGWEVSTVSGRI 288
 DB 457 RHAPNGRPFIDHNTKTTTWEDPRLKIPAHLEKGTSLDTSNDLGLPLPGWEERTHTDGR 516
 QY 289 YFVDHNNRTTQFTDPRLHHMHQCOLKEPSQFLPLPSEGSELEDEELPAQRYERDLVQKL 348
 DB 517 FYINHNKRTOWEDPLENV-----AITGPAVP-----YSRDYKRY 553
 QY 349 KVLHLSLQPOAGHCRIEVSRBEIFPESYRQIKMPPKD-LKKRLMVKTRGEGLDYG 407
 DB 554 EFFRKUKQNDIPNKFEMKLRATVLEDSYRRIMGVRADFLKARLWIEPDGSKLDYG 613
 QY 408 GVAREMLYLLCHEMLNPYYGLFYQYSTDNIYMLQINPDSS1-NPDHLSYFHFVGRIMGLAV 466
 DB 614 GVAREMPFLISKEMENPYGLFEYSATONYTLQINPSGLCNEHDLSYKFEIGRVAGNAV 673
 QY 467 FHGHYNGGFTVPFYKOLLGKPIQLSDLESVDPELHSLVILENDITPVLDTHTFCVHN 526
 DB 674 YHGKLDGFFIRFYQMLHKPITLHDMESVDSEYNSLRWILENDPTE-LDLRFLIDEE 732
 QY 527 AFRILQHELKPNRNPVPTENKKEYVRLYVNRFMGIEBAQFLAQKGNELIPQHL 586
 DB 733 LFGQTHQHELKNGSSEIVVTNKNKXEIYLVITQWRFVNRIOKMAAFKEGFFELIPQDL 792
 QY 587 KPFDKLELILIGLDDKIDLNDKSNTRLKH-CVADSNIVFWQAVETDEERRARLQ 645
 DB 793 KIPDENELELMCLGDDVDNDRHTKYKGYSHQVIOFWFRAVLMDMDSEKRIILQ 852
 QY 646 FVTGSTRVPLQGFALQSGTGAGPRLFTIHLIDANTNLKPAHTCFNRIDIPPYESYK 705
 DB 853 FVTGSTRVPMNGFAELYGSN---GQSFTEQW-GTPEKLPRAHTCFNRLDLPYESFEE 908
 QY 706 LYBKLLTAVEETCGF 720
 DB 909 LWDKQLQWAIENQGF 923

RESULT 10

US-08-539-205A-6

: Sequence 6, Application US/08539205A

: Patent No. 6001619

: GENERAL INFORMATION:

: APPLICANT: Beach, David H.

: APPLICANT: Caligiuri, Maureen

: APPLICANT: Netsky, Bradley

: TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

: NUMBER OF SEQUENCES: 6

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: FOLEY, HOAG & ELIOT LLP

: STREET: One Post Office Square

: CITY: Boston

: STATE: MA

: COUNTRY: USA

ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,205A
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
3-08-539-205A-6

Query Match 31.6%; Score 1228.5; DB 3; Length 834;
Best Local Similarity 40.4%; Pred. No. 8.5e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

93 LSNAISRL-----KDTGYQRDLCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
186 LSEELSRRLQITPDSNGEQFSLIOREPSRLRSCSVTDAVAEQHLPFPFVAVYHTTPG 245
135 IGTGGSV-VDCEGLL-----ENEGTVVDSGGRPLSCFWEPPAPYDSTGAAGGN-- 186
246 LPSGWEERKDAKGRYYVNNRRTTWT-----RPIQLAEDGA-----SGSATNSNNHL 295
187 -----CRFVESPSQORLOAQRLNPDVGRSL-----QTPQNRPHGOSPE----- 227
296 IBPQIRPRSLSPSTVTLAPLEGAKDSPVRRAVKDTLGNPQSPSPYNPKPHQKVTQ 355
228 --LPEGYQRITVQGVYFLHTQGVSTWHDPRIP-----RDLSNVNCDLGLPLPGWEV 280
356 SFLPGWEMWIAPIGRPFIDHNTKTTWEDPRKPFVHMRSKTSLNPNDLGLPLPGWEE 415
281 RTVSGRIYFVDHNNRTQFTDPRLHHIMHQCLKEPSQPLPSEGSLEDEELPAQRY 340
416 RIHLDGRFTFYDHNSKITQWEDPRLOQ-----PAITG-----FAVY 452
698 PYESYKLYEKLTLTAVEETCGF 720

Query Match 31.6%; Score 1228.5; DB 4; Length 834;
Best Local Similarity 40.4%; Pred. No. 8.5e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

93 LSNAISRL-----KDTGYQRDLCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
186 LSEELSRRLQITPDSNGEQFSLIOREPSRLRSCSVTDAVAEQHLPFPFVAVYHTTPG 245
135 IGTGGSV-VDCEGLL-----ENEGTVVDSGGRPLSCFWEPPAPYDSTGAAGGN-- 186
246 LPSGWEERKDAKGRYYVNNRRTTWT-----RPIQLAEDGA-----SGSATNSNNHL 295
187 -----CRFVESPSQORLOAQRLNPDVGRSL-----QTPQNRPHGOSPE----- 227
296 IBPQIRPRSLSPSTVTLAPLEGAKDSPVRRAVKDTLGNPQSPSPYNPKPHQKVTQ 355
228 --LPEGYQRITVQGVYFLHTQGVSTWHDPRIP-----RDLSNVNCDLGLPLPGWEV 280
356 SFLPGWEMWIAPIGRPFIDHNTKTTWEDPRKPFVHMRSKTSLNPNDLGLPLPGWEE 415
281 RTVSGRIYFVDHNNRTQFTDPRLHHIMHQCLKEPSQPLPSEGSLEDEELPAQRY 340
416 RIHLDGRFTFYDHNSKITQWEDPRLOQ-----PAITG-----FAVY 452
698 PYESYKLYEKLTLTAVEETCGF 720

US-09-392-163A-6

RESULT 11
US-09-392-163A-6
Sequence 6, Application US/09392163A
Patent No. 6503742
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Nefsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/392,163A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/539,205
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-392-163A-6

b 453 SREFKQYDYFEKKLKKPADIPNRFEMKLNHNIFESSYRINSKVPDVLKARLMIEFE 512
y 400 GREGLDYGGVAREWYLLCHEWLNPIYGLFOYSTNDNYMLQINDPSI-NPDHLSYFHFV 458
b 513 SBKGLDYGGVAREWFLLSKENFNFPYGLFYSATDNYTLQINPNSGLCNEDHLSYFTEI 572
y 459 GRIMGLAVPHGHYINGGFTVPFYKOLLGKPIQLGSLBSVDPLHKSLLWILENDITPVL 518
b 573 GRVAGLAVPHGLLDGFFIRFPYKMLGKQITLNDMESVDSYNSLKWILENDPTE-LD 631
y 519 HTPCEVHNAGRILOHEUKPNGRNVPTTEENKGYRVLVYVNRFRMGIEAQFALQKGFN 578
b 632 LMFCEDEENFGQTYQDLKPNGBSEIMVTNENKREYDILVIOREFNVRVQKMAFLEGET 691
y 579 ELIPOHLLKPPDKLELILIGLDKIDLNDKSNTRLKHCVADSN-IVRWFQAVETFE 637
b 692 ELLPDLIKIFDENLELLMCGLDVNDVNDROHSIYKNGYCPNHPVIOFWKAVLLMDA 751
y 638 ERARLLQVTVGSTRVPLQGGFKALQGGTGAAGPRLFTHLIDANTMLPKAHTCFNRIDI 697
b 752 EXRIRLLQVTVGSTRVPMNGFAELYGSN---GPQLFTIEQW-GSPEKLPRAHTCFNRIDL 807
y 698 PYVESYKLYEKLTAVERTCQF 720
b 808 PPIETEFEDUREKLLMAVENAQGF 830

RESULT 12
S-08-630-916A-48
Sequence 48, Application US/08630916A
Patent No. 601137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,916A
FILING DATE: 03-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-203
TELEPHONE: (212) 750-9090
TELEFAX: (212) 896-8864/9741
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
IS-08-630-916A-48

Query Match 31.5%; Score 1225; DB 3; Length 906;
Best Local Similarity 35.5%; Pred. No. 2,2e-112;
Matches 281; Conservative 115; Mismatches 227; Indels 169; Gaps 19;

QY 28 PFAKIVDGSQCHSTDTVKNLTDPKWNQHY-----DLVYVKTDSTITISVWNHKKIHKQ 83
DB 188 PNGSALTGDS-QLPFRSDSSGTAVAPE-NRHQPSTNCFGGRS-----RTHRHSG 234
QY 84 AGFLGCVRLLSNAISRLLKDTGYQRLDLCKLNPDSDTDAVRGQIVVSLQTRDR--IGTGGSV 141
DB 235 AS-----ARTTPATGEQSPGARSRRROPVKNSGHS 264
QY 142 VDCGLLENGETVVED-----SGQRPUSCMWEE-----PAPYTDSTGAAAGGN 186
DB 265 GLANGTVNDEPTTATDPEEPSVVGVTSPPAAPLSVTFNPTTSLPAPATPAEG----- 317
QY 187 CRFPESQD--ORLQAOURLNPDVRSLOTPQNRPHGHSPELPE----- 230
DB 318 ---EESTSTQQLPA-----AAQAPALPAGWEQRELPNRGRVYVVDHNTKTTT 363
QY 231 ---GYORTTQQQVYFLHTQTGVSTWHDPRIDRLN----- 264
DB 364 WERPLPQWEKRTDPRGRFYVDHNTRTTTTQRTAAYVYVNRVEQMSQORNQOGAMQHFS 423
QY 265 ---SVNCDLGLPLPGWEVRSTVSGRIYFVDHNNRTTQFTDPR----- 304
DB 424 QRFYQFWSASTDHDPLPGWPKRQD--NGRVYVYVNTNRTTQWEDERTQGMIOEPAL 482
QY 305 ---LHHMNHQCLKEPSQPLPSESGSELEELPAQRYERDLVQKLKVL 352
DB 483 PPGWEMKYTSEGVRYFVDHNTRTTTPKDPGPGESGTKQSGPGAYDRSPRWKYHQFRPLC 542
QY 353 HELSLQPOAGHCHIEVSREIPEESYRQIMKRPDKLKKLNMKFRGEGLDYGGVASE 412
DB 543 HSNAL----PSHVKISVSROTLPFEDSFQIMNMKPYDLRRRLYIIMRGEGLDYGGIAIE 598
QY 413 WLYLLCHEMLNPPYGLFQYSTDNIMLQINPDSSINPDHLSYFVGRIMGLAVFHGYI 472
DB 599 WFFLLSHEVLNPMYCLEYAKKNYCIQINPASSINPDHLYFRFGRPTAMALYHGKFI 658
QY 473 NGGTVVPYKQLGKPKQLSDLSVDDELHKSLLWILENDITPV-LDHTFCVEHNAFGRI 531
DB 659 DTGFTLPFYKRLNKRPTLKDLESIDPEFYNSIVWKENNECEGLYFIQDMEITLKV 718
QY 532 LQHELKNGRNVPTTEENKKEYVRLVYVNRVFMEGIEAQLALQKFNELIPQHLKLPFDQ 591
DB 719 TTHELKEGESIRVTEENKKEYIMLTDRFTGVEEQTKAFLDGNEVAPLEWLVFDE 778
QY 592 KELELIIGLDKIDLNDKSNTRLKHCVADSNIVRWFQAVETTFDEERRARLLQFTVGST 651
DB 779 KELEMLCGMOEIDMSDWKSTIYRHYTKNSKQIQWFWQVVKEMDNKRIQLLQFTVGTG 838
QY 652 RVPLQGGFKALQGGTGAAGPRLFTHLIDANTMLPKAHTCFNRIDIPPYESYKLYEKL 711
DB 839 RLPVGGFAELIGSN---GPQKFCIDKVKET-WLPRSHTCFNRDLDFPYKSYZQLREKLL 894
QY 712 TAVEETCGFAYE 723
DB 895 VAIBETEGFGQE 906

RESULT 13
US-08-630-916A-46
Sequence 46, Application US/08630916A
Patent No. 601137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York

RESULT 14
;-08-247-904B-8
Sequence 8, Application US/08247904B
Patent No. 5981699

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
protein - protein search, using sw model
on on: February 20, 2004, 15:27:02 ; Search time 31.9477 Seconds
(without alignments)
4738.482 Million cell updates/sec
US-10-009-945-2
rfect score: 3884
quence: 1 GSSIKIRLTVLCAKYLAKK.....EKLYEKLITAVEITCGFAVE 723
oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
arched: 801455 seqs, 209382283 residues
otal number of hits satisfying chosen parameters: 801455
nimum DB seq length: 0
ximum DB seq length: 2000000000
st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
tabase : 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pcp.*
10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pcp.*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pcp.*
14: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pcp.*
15: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pcp.*
16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pcp.*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3878	99.8	722	15	US-10-097-534-14
2	3001	77.3	748	12	US-10-021-660-81
3	2913.5	75.0	735	12	US-10-313-955-2
4	2236	57.6	514	12	US-09-764-875-819
5	1471	37.9	766	12	US-10-313-955-4
6	1436.5	37.0	832	12	US-10-032-585-7296
7	1345	34.6	869	15	US-10-128-714-8162
8	1259	32.4	927	15	US-10-097-534-15
9	1249.5	32.2	911	15	US-10-208-823-279
10	1245.5	32.1	739	12	US-10-374-979-89
11	1245.5	32.1	739	15	US-10-097-534-10
12	1245.5	32.1	752	11	US-09-919-039-235
13	1228.5	31.6	834	12	US-10-313-955-6
14	1228.5	31.6	995	15	US-10-097-534-9
15	1228.5	31.6	995	15	US-10-205-823-275

16	1227	31.6	725	15	US-10-185-050-126
17	1226.5	31.6	759	15	US-10-128-714-3162
18	1226	31.6	854	15	US-10-205-823-277
19	1225	31.5	870	15	US-10-097-534-12
20	1225	31.5	906	15	US-10-185-050-48
21	1223	31.5	854	12	US-10-287-218-3
22	1212	31.2	474	11	US-09-774-639-371
23	1212	31.2	474	11	US-09-989-730-249
24	1047	27.0	683	15	US-10-185-050-46
25	1047	27.0	684	15	US-10-097-534-11
26	1023	26.3	380	15	US-10-307-956-32
27	1012	26.1	375	15	US-10-307-956-31
28	1002	25.8	1562	15	US-10-097-534-13
29	836.5	21.5	1094	15	US-10-043-487-300
30	836.5	21.5	1488	15	US-10-043-487-285
31	821.5	21.2	733	15	US-10-097-534-16
32	724	18.6	277	10	US-09-925-300-1628
33	720.5	18.6	358	15	US-10-268-036-5
34	677.5	17.4	276	10	US-09-925-300-1527
35	612	15.8	125	9	US-09-764-870-384
36	612	15.8	125	12	US-09-764-875-1110
37	612	15.8	125	15	US-10-125-340-384
38	518	13.3	823	15	US-10-097-534-25
39	449.5	11.6	1050	15	US-10-097-534-24
40	449.5	11.6	1050	15	US-10-097-534-28
41	449.5	11.6	1054	15	US-10-097-534-29
42	392	10.1	1024	15	US-10-097-534-30
43	389	10.0	584	12	US-10-108-2608-4173
44	389	10.0	1022	12	US-10-287-218-16
45	389	10.0	1022	15	US-10-288-036-2

ALIGNMENTS

RESULT 1
US-10-097-534-14
; Sequence 14, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; TITLE OF INVENTION: MATURATION
; FILE REFERENCE: PLV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/308,958
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-534-14

Query Match 99.8%; Score 3878; DB 15; Length 722;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 722; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 GSSIKIRLTVLCAKYLAKKDFRLPDPFAKIVDVGSGCHSTDTVQNTLDPKMNQHYDLY 61

Db 1 GSSIKIRLTVLCAKYLAKKDFRLPDPFAKIVDVGSGCHSTDTVQNTLDPKMNQHYDLY 60

Qy 62 VGKTSITISVANKHKIKQAGFLGCVRLLSNAISRLKDTGYQRDLCKLNPSDPAV 121

61 VGKTDSTITISVWVHKKHKKQAGAGLGCYVRLLSNALSRLKDTGYORLDCXKLNPSDITDAV 120
122 RGOIVSVLQTRDRIGTGGVWDVCRGLENENGTYYDSGPRPLSCFMEEPAPYTDSTGAA 181
121 RGOIVSVLQTRDRIGTGGVWDVCRGLENENGTYYDSGPRPLSCFMEEPAPYTDSTGAA 180
182 AGGNCRCFVESPSQDORLOARLNPDPVRSGLQTPQNRPHGHQSPELPEGYEORTTVQGO 241
181 AGGNCRCFVESPSQDORLOARLNPDPVRSGLQTPQNRPHGHQSPELPEGYEORTTVQGO 240
242 VYFLHTQTVSTVHWDPRIPRDLSNVCNDBELGPPGWEVSTVSGRIYFVDHNNRTQFT 301
241 VYFLHTQTVSTVHWDPRIPRDLSNVCNDBELGPPGWEVSTVSGRIYFVDHNNRTQFT 300
302 DPLHLMHMHQCOLKEPSOPLPLPSEGSEDELPAPQYERDLVQKLVLRHLSLQOPQ 361
301 DPLHLMHMHQCOLKEPSOPLPLPSEGSEDELPAPQYERDLVQKLVLRHLSLQOPQ 360
362 AGHCRIEVSGREEIFESYRQIMKMRPKDLKKRLMVKFRGEGLDYGGVAREWLYLLCHEM 421
361 AGHCRIEVSGREEIFESYRQIMKMRPKDLKKRLMVKFRGEGLDYGGVAREWLYLLCHEM 420
422 LNPYYGLFOYSTDNIMYLOINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFY 481
421 LNPYYGLFOYSTDNIMYLOINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFY 480
482 KOLLGKPIQLSDLESVDPELHLSLVWILENDITPVLDHTFCVHNFAFRILLOHELKFNKR 541
481 KOLLGKPIQLSDLESVDPELHLSLVWILENDITPVLDHTFCVHNFAFRILLOHELKFNKR 540
542 NVPTEENKEYVRLVNVNRFMRGTEAQFLALOKGFNBLIPOHLLKPPDQKELELIIGGL 601
541 NVPTEENKEYVRLVNVNRFMRGTEAQFLALOKGFNBLIPOHLLKPPDQKELELIIGGL 600
602 DKIDLNDKSNTRLKHCVDADSNIVRFWQAVTFDEERRARILLOFVTGSTRVPLQGFKAL 661
601 DKIDLNDKSNTRLKHCVDADSNIVRFWQAVTFDEERRARILLOFVTGSTRVPLQGFKAL 660
662 QSGTGAAGPRFTIHLIDANTDNLKPAHCTCFNRIDIPPYESYEKLYEKLITAVEETCGFA 721
661 QSGTGAAGPRFTIHLIDANTDNLKPAHCTCFNRIDIPPYESYEKLYEKLITAVEETCGFA 720
722 VE 723
721 VE 722

SULT 2
-10-021-660-81
Sequence 81, Application US/10021660
Publication No. US20030152926A1
GENERAL INFORMATION:
APPLICANT: Murray, Richard
APPLICANT: Glynn, Richard
APPLICANT: Watson, Susan R.
APPLICANT: EOS Biotechnology, Inc.
TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
FILE REFERENCE: 018501-00071005
CURRENT APPLICATION NUMBER: US/10/021,660
CURRENT FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: US/09/784,356
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: US 09/637,977
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 135
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 81
LENGTH: 748
TYPE: PRF
ORGANISM: Homo sapiens
-10-021-660-81

Query Match 77.3%; Score 3001; DB 12; Length 748;
Best Local Similarity 74.6%; Pred. No. 6e-266;
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;
QY 5 IKIRLTVLCAKNLAKKOPFRLPDPFAKIVVDDGSGQCHSTDTVKNLTLPDKNQHLYVVGK 64
DB 12 VKLRLTVLCAKNLAKKOPFRLPDPFAKIVVDDGSGQCHSTDTVKNLTLPDKNQHLYVIGK 71
QY 65 TDSITISVWVHKKHKKQAGAGLGCYVRLLSNALSRLKDTGYORLDCXKLNPSDITDAV 124
DB 72 SDSVITISVWVHKKHKKQAGAGLGCYVRLLSNALSRLKDTGYORLDCXKLNPSDITDAV 131
QY 125 IVVLSQTRDRIGTGGVWDVCRGLENENGTYYDSGPRPLSCFMEEPAPYTDSTGAA 155
DB 132 IVVLSQTRDRIGTGGVWDVCRGLENENGTYYDSGPRPLSCFMEEPAPYTDSTGAA 191
QY 156 --EDSGPRPLSCFMEEPAPYTDSTGAAAGGNCRCFVESPSQDORLOARLNPDPVRSGL 213
DB 192 ASEYSSPGRLSCFVDENTPISTGNTGATCG-----QSSDPRLAERRVRSQRHNYM 242
QY 214 QTPQNRPHGHQSPELPEGYEORTTVQGOVYFLHTQTVSTVHWDPRIPRDLSNVCNDBELG 273
DB 243 ---SRTHLHTPPDLPEGYEORTTVQGOVYFLHTQTVSTVHWDPRIPRDLSNVCNDBELG 298
QY 274 LPPGWEVSTVSGRIYFVDHNNRTQFTDPR-----LHIMNHQCOLKEPSOPLPSEGS 329
DB 299 LPPGWEIRNTATGRVYFVDHNNRTQFTDPRLSANLHLVLRNQKLDQOQQQVW---S 354
QY 330 L---EDELPAQRVERDLVQKLVLRHLSLQOPQOAGHCRIEVSGREEIFESYRQIMKMR 386
DB 355 LCPDTECLTVPRKRDLDVQKLVLRHLSLQOPQOAGHCRIEVSGREEIFESYRQIMKMR 414
QY 387 PKDLKKRLMVKFRGEGLDYGGVAREWLYLLCHEMNPYYGLFOYSTDNIMYLOINPDSS 446
DB 415 PKDLKKRLMVKFRGEGLDYGGVAREWLYLLCHEMNPYYGLFOYSTDNIMYLOINPDSS 474
QY 447 INPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKOLLGKPIQLSDLESVDPELHLSLV 506
DB 475 VNPTEENKEYVRLVNVNRFMRGTEAQFLALOKGFNBLIPOHLLKPPDQKELELIIGGL 534
QY 507 WILENDITPVLDHTFCVHNFAFRILLOHELKPNRNVFVTEENKEYVRLVNVNRFMRGI 566
DB 535 WILENDITPVLDHTFCVHNFAFRILLOHELKPNRNVFVTEENKEYVRLVNVNRFMRGI 594
QY 567 EAQFLALOKGFNEVLPQHLKPPDQKELELIIGGLKIDNDKSNTRLKHCVDADSNIVR 626
DB 595 EAQFLALOKGFNEVLPQHLKPPDQKELELIIGGLKIDNDKSNTRLKHCVDADSNIVR 654
QY 627 NFWQAVETFEERRARILLOFVTGSTRVPLQGFKALQSGTGAAGPRFTIHLIDANTDNL 686
DB 655 NFWQAVETFEERRARILLOFVTGSTRVPLQGFKALQSGTGAAGPRFTIHLIDANTDNL 711
QY 687 KAHTCFNRIDIPPYESYEKLYEKLITAVEETCGFAVE 723
DB 712 KAHTCFNRIDIPPYESYEKLYEKLITAVEETCGFAVE 748

RESULT 3
US-10-313-955-2
; Sequence 2, Application US/10313955
; Publication No. US20030199036A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; Caligiuri, Maureen
; Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA

COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/313,955
FILING DATE: 05-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/392,163
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/539,205
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
S-10-313-955-2

Query Match 75.0%; Score 2913.5; DB 12; Length 735;
Best Local Similarity 72.9%; Pred. No. 5.2e-258;
Matches 552; Conservative 63; Mismatches 71; Indels 71; Gaps 9;
Y 5 IKIRTVLCANLAKDFRFLPDPFAKIVDGSQCCHSTDTVXNTLDPKQNHQYLYVGK 64
b 12 VKRLT-----GLPDPFAKVVVDGSGQCCHSTDTVXNTLDPKQNHQYLYIGK 58
Y 65 TDSITISVNHKTHKKGAGFLGCVRLLSNAISRLXDTGYRLDICKLNPSTDVAVGQ 124
b 59 SDSVTISVNHKTHKKGAGFLGCVRLLSNAISRLXDTGYRLDICKLNPSTDVAVGQ 118
Y 125 IVVSLQTRDRICTGSGVVDRCGLLENE-----GTIVY----- 155
b 119 IVVSLQSRDRICTGSGVVDRCGLFNDLPDGBEERTASGRIQVNLHITFTTOWERPTRP 178
Y 156 --EDSGRPLSCFMEEPAPYTDSTGAAAGGNCRFVESPQQRLOAQRNPDPVRSGL 213
b 179 ASEYSSPGKPLSCFIDENTPISGNTGATCG-----QSSDPRLAERRVRSQRHNTM 229
Y 214 QTPQNRPHGSHOSPPEGYEORTTVQGVYFLHTQTGVSTWHDPRIPRLDINSVNCDELGP 273
b 230 ---SRTHLTPDLPPEGYEORTTVQGVYFLHTQTGVSTWHDPRIPRLDINSVNCDELGP 285
Y 274 LPPGHEVSTVSGRIYFVDHNNRTTQTDPR----LHHIMNHOCQLKEPSQPLPSEGS 329
b 286 LPPGHEVSTVSGRIYFVDHNNRTTQTDPR----LHHIMNHOCQLKEPSQPLPSEGS 341
Y 330 L---EDELPAQRVERDLVQKLVRLHLSLQOPQAGCHRIEVSREEIFEESYRQIMKVRPKDL 386
b 342 LCPDDECTCLTPVRYKRDVQKLVRLHLSLQOPQAGCHRIEVSREEIFEESYRQIMKVR 401
Y 387 PKDLKRLMKFRGEEGLDYGAVAREWLYLLCHEMLNPFYGLFOYSTDNIMYLMQINPDS 446
b 402 PKDLKRLMKFRGEEGLDYGAVAREWLYLLCHEMLNPFYGLFOYSTDNIMYLMQINPDS 461
Y 447 INPDHLSYFHVGRIMGLAVFHGHYINGGTFVFPYKQLCKPIQLSDLESVDPDLHSLV 506
b 462 VNPHLSYFHVGRIMGLAVFHGHYINGGTFVFPYKQLCKPIQLSDLESVDPDLHSLV 521
Y 507 WLENDITPVLDHTFCVEHNAGFRILQHELKPNGRNVPVTEENKKEYVRLYVNNRFLRGI 566

Db 522 WLENDITGVLDHTFCVEHNAYGEIIQHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGI 581
QY 567 BAQFLAQKGFNELIPOHLLKPFDOKELELIIGLDKIDLNWDKSNTRLKHCVCVADSNIVR 626
Db 582 BAQFLAQKGFNEVPOHLLKPFDOKELELIIGLDKIDVNWDKSNTRLKHCCTPDSNIVK 641
QY 627 WFOAVETFOBERRARRLQFVTGSTRVPLQGFKALQSGTAGAPRLFTTHLIDANTDNL 686
Db 642 WFKAVAFFDEERRARLLQFVTGSSRPVQGFKALQ---GAAGPRLFTTHQIDACTNNLP 698
QY 687 KAHTCFNRIDIPPYESYEKLYEKLTAVEETCGFAVE 723
Db 699 KAHTCFNRIDIPPYESYEKLYEKLTAIBETCGFAVE 735
RESULT 4
US-09-764-875-819
; Sequence 819, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 819
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-819

Query Match 57.6%; Score 2236; DB 12; Length 514;
Best Local Similarity 80.7%; Pred. No. 5.2e-196;
Matches 414; Conservative 43; Mismatches 42; Indels 14; Gaps 4;
QY 218 NRPHGSHOSPPEGYEORTTVQGVYFLHTQTGVSTWHDPRIPRLDINSVNCDELGP 277
Db 9 SRTHLTPDLPPEGYEORTTVQGVYFLHTQTGVSTWHDPRIPRLDINSVNCDELGP 68
QY 278 WEVRASTVSGRIYFVDHNNRTTQTDPR---LHHIMNHOCQLKEPSQPLPSEGS 330
Db 69 WEVRASTVSGRIYFVDHNNRTTQTDPR---LHHIMNHOCQLKEPSQPLPSEGS 124
QY 331 EDELPAQRVERDLVQKLVRLHLSLQOPQAGCHRIEVSREEIFEESYRQIMKVRPKDL 390
Db 125 DTECLTPVRYKRDVQKLVRLHLSLQOPQAGCHRIEVSREEIFEESYRQIMKVRPKDL 184
QY 391 KCRLMKFRGEEGLDYGAVAREWLYLLCHEMLNPFYGLFOYSTDNIMYLMQINPDS 450
Db 185 WKRLMKFRGEEGLDYGAVAREWLYLLCHEMLNPFYGLFOYSTDNIMYLMQINPDS 244
QY 451 HLSYFHVGRIMGLAVFHGHYINGGTFVFPYKQLCKPIQLSDLESVDPDLHSLV 510
Db 245 HLSYFHVGRIMGLAVFHGHYINGGTFVFPYKQLCKPIQLSDLESVDPDLHSLV 304
QY 511 NDITPVLDHTFCVEHNAGFRILQHELKPNGRNVPVTEENKKEYVRLYVNNRFLRGI 570
Db 305 NDITPVLDHTFCVEHNAGFRILQHELKPNGRNVPVTEENKKEYVRLYVNNRFLRGI 364
QY 571 LALQKGFNELIPOHLLKPFDOKELELIIGLDKIDLNWDKSNTRLKHCVCVADSNIVR 630
Db 365 LALQKGFNEVPOHLLKPFDOKELELIIGLDKIDVNWDKSNTRLKHCCTPDSNIVK 424
QY 631 AVTFOBERRARRLQFVTGSTRVPLQGFKALQSGTAGAPRLFTTHLIDANTDNL 690
Db 425 AVEFFDEERRARLLQFVTGSSRPVQGFKALQ---GAAGPRLFTTHQIDACTNNLP 481
QY 691 CFNRIDIPPYESYEKLYEKLTAVEETCGFAVE 723
Db 482 CFNRIDIPPYESYEKLYEKLTAIBETCGFAVE 514

SULT 5

-10-313-955-4

Sequence 4, Application US/10313955

Publication No. US20030199036A1

GENERAL INFORMATION:

APPLICANT: Beach, David H.

Caigiuri, Maureen

Nesley, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/392,163

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

FILING DATE: 05-Dec-2002

297 RYTPGPRYFVDHNRITTTTWDPRRQYIIRSYGPNNTIQQQPVSQLGELPSPGWEMLT 356
284 VSGRIYFVDHNRITTTQFDPRLHHIMNHQCOLKEPSQPLPSEGLSEDEELPAQRYERD 343
357 NTARVYFVDHNRITTTTWDPR-----LPS--SL-DQNVF--QYKRD 392
344 LVOKLVRLHSLSLQOQAGHCRLEVSREBIEFESYPOIMKMRPKOLKELMVKESGEG 403
393 FRKLIYFLSOPAL-HPLPGOCHI KVRNHFEDSYABIMRQATDLDKRLMKFDGEDG 451
404 LDYGVAREMILYLLCHEMLNPPYGLFOYSTDNIMYMLQINPDSSINPDHLSYFFHVGRI 463
452 LDYGLSREYFLLSHEMFNPFYCLFEYSSVDNYTLQINPHSGINPEHLNPFYKFGVIG 511
464 LAVFHGVIYNGFTVPFYKQLLGPQLSLESVDPELHKLWILENDITPVLDTHTFCV 523
512 LAIFRRFVDAPFVVSFYKMLQKVTLOQKESMDAIEYYSRLWILDNDITGVLDLTFV 571
524 EHNAPGRILQHELKPNGRNVPTVENKKEYVRLVNNRFRMRGIEAQFLALQKGFNELIPQ 583
572 EDCGGEVVITDLKENGNIETVENKKEYVVDLVTVW-IQKRIEEQFNAPHEGFSSELIPO 630
584 HLLKPFQKLELIIIGLDKIDLDNDWNSNTRKHCVADSNIVRWFOAVETFEERRARL 643
631 ELINVFDERELELLIGISIDMEDWKKHDKYRSYSENDQIIKWFELMDWESNEKKSRL 690
644 LOFTVGTSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTONLPAKTCFNRIDIPPPYESY 703
691 LOFTVGTSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTONLPAKTCFNRIDIPPPYESY 746
704 EKLYEKLITAVEETCGFAVE 723
747 KOLDHKLSTAVEETIGFQOE 766

RESULT 6
US-10-032-585-7296
; Sequence 7296, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7296
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7296

Query Match 37.0%; Score 1436.5; DB 12; Length 832;
Best Local Similarity 38.4%; Pred. No. 2.6e-122;
Matches 326; Conservative 124; Mismatches 233; Indels 167; Gaps 21;

7 IRLTVLCANLAKDPFRLPDPFAKIVVDGSGQCHSTDTVKNTLDPKWNQHYDLYVGKTD 66
17 INVKVAESLYKRDVFPQDPFAVLTVDSG-QTKTITAKTLNPNYNETFN-FOAKED 74
67 SI-TTSVWNHKKIHKQAGFLGCVRL-LSNAISRLKDTGYQRL--DLCKLNPSDTPVR 122
75 SILVQVDFDQKX-KKKQDQGLGVINVRIGVIDLSNSESSEITITDLKSN--ENLAVS 131
123 GOIVVSLQ-----TRD-----RIGTGGSVVDGRLLE 149
132 GKLIIVISHNRNSGGVYTTATTGTGASSNNIATITSGVNNLRIGSATNTANSTAQA 191

64 SIKRLTVLCANLAKDPFRLPDPFAKIVVDGSGQCHSTDTVKNTLDPKWNQHYDLYVG 63
7 SRRIRVIAADGLYKRDVFRFPDPFAVLTVDSG-QTKTITAKTLNPNYNETFN 65
64 KTDSTISVWNHKKIHKQAGFLGCVRLSNAISRLKDTGYQRL--DLCKLNPSDTPVR 121
66 DNSTIALQVDFDQKX-KKKQDQGLGVINVRIGVIDLSNSESSEITITDLKSN--ENTV 122
122 RQIVVSLQ-----TRDRI GTGGSVVDGRLLENETGYVED 157
123 HGKIIINLSITTAQSTLQVPSSAAGARTQRTSITNDPQSSKSSSVRNPASSRAGSPTRD 182
158 SGPG-----RPLSCFME-----PAPY-----TDSTG-----AAA 182
183 NAPAASPASSEPTFSFEDQYGRLPFGWERTDNLGRYYVDHNRITTTWIRPNLSVA 242
183 GGGNCRFVESPS-----QDORLOAQLRNPDRVSGSLQTPQNRPHGQSPPELPEGYEQ 234
243 GAAAAELHSSASSANVTGVPSSNAARTEASVLTSNATTAG-----SGELPPGWEQ 296
235 RTTVQGVYFLHTQTGVSTVHWDPRIDRLNSVN-----CDELGPLPGWEVRST 283

Query Match 37.9%; Score 1471; DB 12; Length 766;
Best Local Similarity 40.9%; Pred. No. 1.6e-125;
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
-10-313-955-4

37.9%; Score 1471; DB 12; Length 766;
Best Local Similarity 40.9%; Pred. No. 1.6e-125;
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;

725 DLPPKSYEQLEKELLPALTEETGQCE 752

ULT 13

10-313-955-6

sequence 6, Application US/10313955

publication NO. US20030199036A1

GENERAL INFORMATION:

APPLICANT: Beach, David H.

Caligiuri, Maureen

Nefsky, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/313,955

FILING DATE: 05-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/392,163

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/539,205

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: CSV-005.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 834 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

10-313-955-6

Query Match 31.6%; Score 1228.5; DB 12; Length 834;
 Best Local Similarity 40.4%; Pred. No. 3.2e-103;
 Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;
 93 LSNALSRLL-----KDTGYQRLLDCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
 186 LSELSRLQLITPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQGLPPPSVAYVHTTPG 245
 135 IGTGGSV-VDCRGLL-----ENEGTVYEDSGPGRPLSCFMEEAPYTDSTGAAGGN-- 186
 246 LPSCWEERKDAKGRYYVNNHNNRTTWT-----RPIQLAEDGA-----SGSATNSNNHL 295
 187 -----CRFVESPSQDQRLQALRNPDRVGLS-----QTPQNRPHGHQSFE----- 227
 296 LEPQIRRRPSRLSSPTVTLTSLAPLEGAKDSPVRAVKDTLSNPQSPQSPYNSPKPQKVQ 355
 228 --LPGYEQRTVQGVVFLTQTGVSTWHDPRIP-----RDLNSVNCDELGLPLPGWEV 280
 356 SFLPPGWMRIAPNGRPFFIDHNTKTITWEDPRLKFVFMRSKTSLSNPDLGLPLPGWEE 415
 281 RSTVSGRIYFVDHNNRTTQFTDPRLLHIMNHQCCQLKEPSPQLPLPSGCSLEDELPQRY 340
 416 RHLDGRFYIDHNSKIQTWEDPRLON-----PAITG-----PAPVY 452

QY 341 ERDLVQKLVLRHLSLQOPQAGHCRIEVSREEIFESYSYQIMK- RPKOLKKELMYKER 399
 DB 453 SREFKQKDYDFRKKLKPADIPNREFMKLHNNIFESSYRIMSVKPPDVLKARLWIEFE 512
 QY 400 GEGLDYGGVAREWLYLLCHEMLNPYYGLFOYSTDNIMYLQINPDSSI-NPDHLSYFHFV 458
 DB 513 SEKGLDYGGAAREWFFLLSKEMNFYYGLFEYSATDNYTLQINPNSGLCNEDHLSYFTFI 572
 QY 459 GRIMGLAVFHGHYINGGTFPPYKQLLGKPIQLSDLESVDPELHKSLSVWILENDITPVL 518
 DB 573 GRVAGLAVFHGKLLDGFIRPFPYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 631
 QY 519 HTPFCVHNAGFRILQHELKPNGRNVPTTEENKCEYVRLYVNNRFRMGIEAQFLALQKGF 578
 DB 632 LMFCIDEENFGQTYQVLDKPNKNGSEIMVTNENKREVIDLVIQRFVNRVQKQMNAPLEGFT 691
 QY 579 ELIPQHLLKFPQKLELELIIGGLDKIDLNDKNSNTRLKHCVADSN-IVRWFWQAVEIFDE 637
 DB 692 ELLPDLILKIFDENSELELLMCGLDGVNDVRQHSIYKNGYCPNHPVQWFWKAVLLMDA 751
 QY 638 ERRARLLQFVGTSTVPLOGFKALOGSTGAAGPRLTTHLIDANTDNLPKAHTCFNRIDI 697
 DB 752 EXRIQLQFVTGTSRVFNGFAELYGSN---GPQLFTIEQW-GSPEKLPRAHTCFNRDL 807
 QY 698 PYESYEKLYEKLITAVEETQGF 720
 DB 808 PYTEFEDLRKLLMAVENAQGF 830

RESULT 14

US-10-097-534-9

; Sequence 9, Application US/10097534

; Publication No. US20030049607A1

; GENERAL INFORMATION:

; APPLICANT: GREENER, TSVIKA

; APPLICANT: MOSKOWITZ, HAIM

; APPLICANT: REISS, YUVAL

; APPLICANT: AUKOY, IRIS

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL

; TITLE OF INVENTION: MATURATION

; FILE REFERENCE: PLV-001.01

; CURRENT APPLICATION NUMBER: US/10/097,534

; PRIOR FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/275,224

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: 60/308,958

; PRIOR FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: 60/340,170

; PRIOR FILING DATE: 2001-12-07

; NUMBER OF SEQ ID NOS: 71

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 9

; LENGTH: 995

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-097-534-9

Query Match 31.6%; Score 1228.5; DB 15; Length 995;
 Best Local Similarity 40.4%; Pred. No. 4.3e-103;
 Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;
 93 LSNALSRLL-----KDTGYQRLLDCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
 347 LSELSRLQLITPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQGLPPPSVAYVHTTPG 406
 QY 135 IGTGGSV-VDCRGLL-----ENEGTVYEDSGPGRPLSCFMEEAPYTDSTGAAGGN-- 186
 DB 407 LPSCWEERKDAKGRYYVNNHNNRTTWT-----RPIQLAEDGA-----SGSATNSNNHL 456
 QY 187 -----CRFVESPSQDQRLQALRNPDRVGLS-----QTPQNRPHGHQSFE----- 227
 DB 457 LEPQIRRRPSRLSSPTVTLTSLAPLEGAKDSPVRAVKDTLSNPQSPQSPYNSPKPQKVQ 516

```
228 --LPEGYEORTTVQGVFLHTQTGVSTWHDPRIP-----RDLSVNCDELGLPLPGWEV 280
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 SFLPPGWNRIAPNGRPFFIDHTTWTWEDPRKLPVHMRKSTSLNPNLGLPLPGWEE 576
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 RSTVSGRIYFVDHNNRTTQTPRLHHIMHQCQLKPSQPLPLPSEGSLEDELPQRY 340
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
577 RIHLDRGRTFYIDHNSKITQWEDPRLQN-----PAITG-----PAVPY 613

341 ERDLVOKLVLRHLSLQOQAGHCRIEVSREIFESYRQIMKM-RPKDLKRLMKVFR 399
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
614 SREFKQKYDYFRKKLKPADIPIRNFEMKLHNNIFESYRIRMSVKRPOVLKARLWIEFE 673
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
400 GEGLDYGGVAREWMLLCHMLNPYYGLFOYSTDNITMLOINPDSSI-NPDHLSYHFV 458
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
674 SEGGLDYGGVAREWFLLSKEMFNYYGLFEYSATDNTYTLQINPNSGLCNEDHLSYTFI 733
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
459 GRIMGLAVFHGHIYNGGFTVPFYKOLLGKPIQLSDLESVDPELHKSILWILENDITPVLD 518
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
734 GRVAGLAVFHGKLDGFFIRFPFYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 792
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
519 HTFCVEHNAFGRILQHELKPNRNVPTVENKKEYRVLVYVWRFMRGIEAQLALQKGFN 578
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
793 LMFCDIENFGQTYQVDLKNPNSBIMVTNENKREYIDLVIQWRFVNRVQKQMAFLEGFT 852
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
579 ELIPQHLKPFDOKELELIIGGLDKIDLNDWKSNTLKHCVADSN-IVRFWQAVETFE 637
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
853 ELLPIDLIKIFDENELELMCGLDGVDVNDWRQHSIYKNGYCPNHPVIOQFWKAVLLMDA 912
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
638 ERARLLOQVGTSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTDNLKHAHTCFNRIDI 697
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
913 EKRIELLOQVGTSTRVPMNGFAELVGSN---GPQLFTIEQW-GSPEKLPRACHTCFNRDL 968
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
698 PYESYEKLYEKLITAVEETCGF 720
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
969 PPYETFEFLREKLMAVENAQGF 991
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

sequence 275, Application US/10205823
Application No. US20030108963A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Wensey, Angela M.
APPLICANT: Glatc, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
EQ ID NO 275
LENGTH: 995
```

```
      ; TYPE: PRT
      ; ORGANISM: Homo sapiens
      ; US-10-205-823-275

      Query Match      31.6%; Score 1228.5; DB 15; Length 995;
      Best Local Similarity 40.4%; Pred. No. 4.3e-103;
      Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

      QY 93 LSNAISRL-----KDTGYQRLDLCKLNPDS-----TDVRCQ-----IVVSLQTRDR 134
      DB 347 LSEELSRQLQITPDSNGEQFSSLIQREPSRLRSCVSTDAVAEQGHLPPSPVAYVITPG 406
      QY 135 ICTGGSV-VDCRGLL-----ENEGTVYEDSGPRPLSCFMEEPAPYTDSTGAAAGCN-- 186
      DB 407 LPSGWEERKDAKGRYYVNNHNRITTTWT-----RPIQLAEDGA-----SGSATNHNHL 456
      QY 187 -----CRFVESQSQRLOAQLRNPVDYRGL-----QTPQNRPHCHQSPE----- 227
      DB 457 IEPQIRRPRLSSPTVTLSPLEGAKDSPVERAVKDTLSNPQSPQSPSPYNSPKQHKVTQ 516
      QY 228 --LPEGYEORTTVQGVFLHTQTGVSTWHDPRIP-----RDLSVNCDELGLPLPGWEV 280
      DB 517 SFLPPGWNRIAPNGRPFFIDHTTWTWEDPRKLPVHMRKSTSLNPNLGLPLPGWEE 576
      QY 281 RSTVSGRIYFVDHNNRTTQTPRLHHIMHQCQLKPSQPLPLPSEGSLEDELPQRY 340
      DB 577 RIHLDRGRTFYIDHNSKITQWEDPRLQN-----PAITG-----PAVPY 613
      QY 341 ERDLVOKLVLRHLSLQOQAGHCRIEVSREIFESYRQIMKM-RPKDLKRLMKVFR 399
      DB 614 SREFKQKYDYFRKKLKPADIPIRNFEMKLHNNIFESYRIRMSVKRPOVLKARLWIEFE 673
      QY 400 GEGLDYGGVAREWMLLCHMLNPYYGLFOYSTDNITMLOINPDSSI-NPDHLSYHFV 458
      DB 674 SEGGLDYGGVAREWFLLSKEMFNYYGLFEYSATDNTYTLQINPNSGLCNEDHLSYTFI 733
      QY 459 GRIMGLAVFHGHIYNGGFTVPFYKOLLGKPIQLSDLESVDPELHKSILWILENDITPVLD 518
      DB 734 GRVAGLAVFHGKLDGFFIRFPFYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 792
      QY 519 HTFCVEHNAFGRILQHELKPNRNVPTVENKKEYRVLVYVWRFMRGIEAQLALQKGFN 578
      DB 793 LMFCDIENFGQTYQVDLKNPNSBIMVTNENKREYIDLVIQWRFVNRVQKQMAFLEGFT 852
      QY 579 ELIPQHLKPFDOKELELIIGGLDKIDLNDWKSNTLKHCVADSN-IVRFWQAVETFE 637
      DB 853 ELLPIDLIKIFDENELELMCGLDGVDVNDWRQHSIYKNGYCPNHPVIOQFWKAVLLMDA 912
      QY 638 ERARLLOQVGTSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTDNLKHAHTCFNRIDI 697
      DB 913 EKRIELLOQVGTSTRVPMNGFAELVGSN---GPQLFTIEQW-GSPEKLPRACHTCFNRDL 968
      QY 698 PYESYEKLYEKLITAVEETCGF 720
      DB 969 PPYETFEFLREKLMAVENAQGF 991
```

Search completed: February 20, 2004, 15:37:43
Job time : 35.9477 secs

66 DNSTIAIQVFDQKF-KKGGQGLGVINLRVGVLDLAIGGDEMLTRDLKSN--ENTVV 122
122 RGQIVVSLQ-----RPLSCFMBE---PAPY---TDSTG-----AAA 182
123 HGKLIINLSTTAQSTLQVPSAASGARTQRTSITNDPQSSQSSSVSRNPASSRAGSPTRD 182
158 SGPG-----RPLSCFMBE---PAPY---TDSTG-----AAA 182
183 NAPAASPASSERPTFSFEDQYGLPFGWERRTDNLGRYYVDHNRSTTWIRPNLSSVA 242
183 GCGNCRFVESPS-----QDRLQALRNPDVRSLOTPQNRPHGHSQSPBELPGEVQ 234
243 GAAAEHLSSASSANVTGVPSSSNAARTEASVLTSTNATTAG-----SGLPFGHEQ 296
235 RTVQGVYFLHTQTVSTWDPRIPDLNSVN-----CDELGPLPGEVVRST 283
297 RYTPGGRFYFDHNRTRTTTWDPRQYIRSYGGPNNAITQQQPVSQLGFLPSGHEMRLT 356
284 VSGRIYFVDHNRTRTTQFTDPRLHHMNHQCOLKEPSQPLPSEGSGLEDEELPAQRYERD 343
357 NTARYFVDHNRTRTTTWDPR-----LPS--SL--DQNVF--QYKRD 392
344 LVQKLVLRLHLSLQOPAL-HPLPGQCHIKVRNHI FDSYAEIMRQSATDLKRLMIKFDGEDG 451
393 FRKLIYLSQAL-HPLPGQCHIKVRNHI FDSYAEIMRQSATDLKRLMIKFDGEDG 451
404 LDYGVAREWLYLLCHEMLNPPYGLFOYSTDNIMYQINPDSSINPDHLSYFHFVGRIMG 463
452 LDYGLSREYFLLSHEMFNPFYCLFEYSVDNYTLQINPHSGINPEHLNFKFTGRVIG 511
464 LAVFHGHVINGGFTVPFYKOLLGKPIQLSDLESVDPELHKLVLWILENDITPVLDTTCV 523
512 LAIFHRRFVDAPFVVSFYKMLQKVTLDQMSMDAAYRSLVWILDNDITGVLDLTFV 571
524 ENAFGRILQHELKPNRNVPTVEENKEVYLYVWRFMRGTEAQFLALQKGFNELIPQ 583
572 EDCNCFGEVVTIDLPKNGRNIETVEENKEVYLYVWRFMRGTEAQFLALQKGFNELIPQ 631
584 HLLKFPDQKELELIGGLDKDLDNDKSNRLKHCVDASNIVRWFQAVETFDERRARL 643
632 ELINVFDERELELLIGGLSEIDMEDWKHKDYRSYENDQIIKFWELMDENSKKSL 691
644 LQFVTGSTRVPLQGFKALQGSTGAGPRLETHILIDANTDNLKHAHTCFNRDIPPEY 703
692 LQFTTGTSTRIFVNGFKDLQSGD---GPRKFTIEKA-GEPNKLPKHAHTCFNRDLDPYTSK 747
704 EKLVEKLTAVEETCGFAVE 723
748 KOLDHKLSTIAVEETIGFQGE 766

RESULT 2
ubiquitin-protein ligase (EC 3.2.19) - fission yeast (Schizosaccharomyces pombe)
;Alternate names: E6-AP-like protein ubiquitin ligase
;Species: Schizosaccharomyces pombe
;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
;Accession: S66562; T45159
;RefSeq: B.; Beach, D.
;Title: Pubi acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25
;Reference number: S66562; MUID:96205868; PMID:8635463
;Accession: S66562
;Status: Preliminary; nucleic acid sequence not shown
;Molecule type: mRNA
;Residues: 1-766 <NER>
;Cross-references: GB:U66716; NID:g1519443; PID:AAS07514.1; PID:g1519444
;RefSeq: B.S.; Beach, D.
;Description: Pubi acts as an E6-AP-like protein ubiquitin ligase in the degradation of
;Reference number: 222935
;Accession: T45159
;Status: Preliminary; translated from GB/EMBL/DBD

A:Molecule type: mRNA
A:Residues: 1-766 <NER>
A:Cross-references: EMBL:Y07592; PIDN:CAA68867.1
C:Genetics:
A:Gene: pubi
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C:Keywords: ligase
F:205-242/Domain: WW repeat homology <WW1>
F:288-325/Domain: WW repeat homology <WW2>
F:345-382/Domain: WW repeat homology <WW3>

Query Match 37.9%; Score 1471; DB 1; Length 766;
Best Local Similarity 40.9%; Pred. No. 6.3e-95; Indels 120; Gaps 20;
Matches 327; Conservative 118; Mismatches 33;

QY 4 SIKRLTVLCARNLAKKDFRLLPDPFAKIVVDSGGCHSTDTYKNTLDPKMNQHYDLYVG 63
DB 7 SRIRVTIVAADGLYKRDVFRFPDPFAVLTVDGB-QTHTTTAKKTLNPNVNETPEVNT 65
QY 64 KTDSTISVMNHKIKKQAGFLGCVRLLSNLSIKDTGYQL--DLCKLNPSTDVA 121
DB 66 DNSTIAIQVFDQKF-KKGGQGLGVINLRVGVLDLAIGGDEMLTRDLKSN--ENTVV 122
QY 122 RGQIVVSLQ-----RPLSCFMBE---PAPY---TDSTG-----AAA 182
DB 123 HGKLIINLSTTAQSTLQVPSAASGARTQRTSITNDPQSSQSSSVSRNPASSRAGSPTRD 182
QY 158 SGPG-----RPLSCFMBE---PAPY---TDSTG-----AAA 182
DB 183 NAPAASPASSERPTFSFEDQYGLPFGWERRTDNLGRYYVDHNRSTTWIRPNLSSVA 242
QY 183 GCGNCRFVESPS-----QDRLQALRNPDVRSLOTPQNRPHGHSQSPBELPGEVQ 234
DB 243 GAAAEHLSSASSANVTGVPSSSNAARTEASVLTSTNATTAG-----SGLPFGHEQ 296
QY 235 RTVQGVYFLHTQTVSTWDPRIPDLNSVN-----CDELGPLPGEVVRST 283
DB 297 RYTPGGRFYFDHNRTRTTTWDPRQYIRSYGGPNNAITQQQPVSQLGFLPSGHEMRLT 356
QY 284 VSGRIYFVDHNRTRTTQFTDPRLHHMNHQCOLKEPSQPLPSEGSGLEDEELPAQRYERD 343
DB 357 NTARYFVDHNRTRTTTWDPR-----LPS--SL--DQNVF--QYKRD 392
QY 344 LVQKLVLRLHLSLQOPAL-HPLPGQCHIKVRNHI FDSYAEIMRQSATDLKRLMIKFDGEDG 451
DB 393 FRKLIYLSQAL-HPLPGQCHIKVRNHI FDSYAEIMRQSATDLKRLMIKFDGEDG 451
QY 404 LDYGVAREWLYLLCHEMLNPPYGLFOYSTDNIMYQINPDSSINPDHLSYFHFVGRIMG 463
DB 452 LDYGLSREYFLLSHEMFNPFYCLFEYSVDNYTLQINPHSGINPEHLNFKFTGRVIG 511
QY 464 LAVFHGHVINGGFTVPFYKOLLGKPIQLSDLESVDPELHKLVLWILENDITPVLDTTCV 523
DB 512 LAIFHRRFVDAPFVVSFYKMLQKVTLDQMSMDAAYRSLVWILDNDITGVLDLTFV 571
QY 524 ENAFGRILQHELKPNRNVPTVEENKEVYLYVWRFMRGTEAQFLALQKGFNELIPQ 583
DB 572 EDCNCFGEVVTIDLPKNGRNIETVEENKEVYLYVWRFMRGTEAQFLALQKGFNELIPQ 630
QY 584 HLLKFPDQKELELIGGLDKDLDNDKSNRLKHCVDASNIVRWFQAVETFDERRARL 643
DB 631 ELINVFDERELELLIGGLSEIDMEDWKHKDYRSYENDQIIKFWELMDENSKKSL 690
QY 644 LQFVTGSTRVPLQGFKALQGSTGAGPRLETHILIDANTDNLKHAHTCFNRDIPPEY 703
DB 691 LQFTTGTSTRIFVNGFKDLQSGD---GPRKFTIEKA-GEPNKLPKHAHTCFNRDLDPYTSK 746
QY 704 EKLVEKLTAVEETCGFAVE 723
DB 747 KOLDHKLSTIAVEETIGFQGE 766

RESULT 3

9744 obable ubiquitin-protein ligase [imported] - Neurospora crassa
Alternate names: protein B24B19.160
Species: Neurospora crassa
Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
Accession: T49744
Schulte, U.; Algn. V.; Kohseisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, R.
Submitted to the Protein Sequence Database, May 2000
Reference number: Z25022
Accession: T49744
Status: preliminary
Molecule type: DNA
Residues: 1-815 <SCH>
Cross-references: EMBL:AL356192; GSPDB:GN00116; NCSP:B24B19.160
Experimental source: BAC clone B24B19; strain OR74a
Genetics:
Gene: NCSP:B24B19.160
Map position: 6
Introns: 11/1; 24/1; 59/2; 110/1; 783/2
Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
219-276/Domain: WW repeat homology <WW1>
334-371/Domain: WW repeat homology <WW2>
393-430/Domain: WW repeat homology <WW3>
Query Match 37.0%; Score 1437; DB 2; Length 815;
Best Local Similarity 38.3%; Pred. No. 2.1e-96;
Matches 321; Conservative 125; Mismatches 238; Indels 154; Gaps 21;
9 LTVLCAKNIKKAKKDFRLPPFAKIVDGGSGCHSTDTVQNTLDPKWNHGYDLV-----V 62
9 LPVIAADGLYKRDVFRPPFPFVATNGE-QTKTVQSKRTLPNTWNEHDFWLIQSKV 67
63 GKTDSTITISVNNHKKHKOGAGFLGCVRL-LSNAISRLKDTGYQRLCLKNPDSPTDAV 121
68 NEDSILLAVQVFDQKPK-KKXQDQGLGVINVRIGDVIELAPDAEDQMLRDLKSKSDNLV 126
122 RGQIVSLQTR-----DRITGSGVDCRGLLENCTVYED-----SGP----- 160
127 HGKLIINLSNLTATMSRLGFPSSRRPLLTFQSSVINDRANERPSSMSGPNGTANN 186
161 ----GRPLSCFMEE-----PA-----PYTDSGT----- 179
187 MTLASRPASLAVSSSTAPPTGNGTAPNTPLVPAQARHSHSTLSPFSDSGRLPAGWE 246
180 -----AAAGGNCRFVSPSQDORLOAQRLENDVDSGLQT 215
247 RREDHLGRTYVDHNSRTTSWNRPTGTGAENRTABANTQVERQHRNLTLPEDRTGANS 306
216 PQRNPH-----GHOSP---ELPEGYEORTVQGVYFLHTQTGVSTWHDPR 258
307 PTLOQQAATAANATMMHTGATPTGELPAGWEQRFTEGEPYFVDHNTTITTTWVDP 366
259 IPRDL-----NSVN-----CDELGLPPGMEVSTVSGRIYFVDHNNRTQTQDPR 305
367 RQOIVRYMGONTNGTITQQQVSQLGPLSPGWEMLTNTARVYFVDHNTKTITWDDPR 425
306 HHINWQCQLKEPQPLPSEGLSEDEELPAQYERDLVKVLRHLSLQCPQAGHC 365
426 -----LPS--SL--DQNVF--QYKDRFRLKIIFRSPAV-RWMSGQC 461
366 RIEVSREEIEEYSRQIMKRPKDLKKRLMKVPRGEGLDYGVGAEMWLYLCHENLNPY 425
462 HIKVRRSHIPEDSFAEISRQSAUTDKRLMIKTDGSDGLDYGLSREFFFLLSHEMFNP 521
426 YGLQYSTDNTYMLQINDSPINDPHLSYFHFVGRINGLAVFHHVINGGFTVYFKQL 485
522 YCLFEYSANDNTYQINPHSGINPEHLNYPFKFIVGVVGLAIFRRFLDAFFGALYKMWL 581
486 GKPIQLSPLESVDPELHKSLWILENDITVLDHTFCVEHNAFGRLOHELKPNGRNVVP 545
582 GKAVSLADMEGVADAFHRSLOWMLNDITDVLDTATSTEDEREGVITEEDLIPNGRNIAV 641
546 TEENKKEVRLVNWVRFWRGIEAQFLAQGFNELIPQHLKPFDDQKELELIGGLDKID 605

Db 642 TNEKKEVELMWVRTEKLEQCFRAKDFGFLHLPQDLINVFDERELLELGIAEID 701
QY 606 LNDWKSNTRLKXCHVADSNVWFMQAVETDEERRARLLQFTVGSTRVLPQKALQGST 665
Db 702 VDDMKKHTDYRGYTESDEVIOFQWQTVRSWDGSKSLQLOFTTGTSRIPVNGFKDJQSGSD 761
QY 666 GAAGPRLTTHILIDANTDNLPKATCFNRIDIPYESYEKLKLTAVETCGFAVE 723
Db 762 ---GPRRTIEKAGEIT-NLPKATCFNRDLDPYKSLMLQKLTIAVETWGFQGE 815
RESULT 4
S43217
ubiquitin-protein ligase (EC 6.3.2.19) RSP5 - yeast (Saccharomyces cerevisiae)
N;Alternate names: E6-AP-like protein ubiquitin ligase; protein YER125w; PUB1 protein
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C;Accession: S43217; S50628; S70050
R;Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor
submitted to the EMBL Data Library, February 1993
A;Reference number: S30812
A;Accession: S43217
A;Molecule type: DNA
A;Residues: 1-809 <MUL>
A;Cross-references: GB:U18916; EMBL:L11119; NID:G1384128; PID:AC03223.1; PID:G603364
R;Dietrich, F.S.
Submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, and lambda
A;Accession: S50628
A;Molecule type: DNA
A;Residues: 1-809 <DIE>
A;Cross-references: EMBL:U18916; NID:G1384128; PID:AC03223.1; PID:G603364; GSPDB:GN0
Mol. Microbiol. 18, 77-87, 1995
A;Title: NPI1, an essential yeast gene involved in induced degradation of Gap1 and Fur
A;Reference number: S70050; MUID:96154942; PMID:8596462
A;Accession: S70050
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-101 <HEI>
C;Genetics:
A;Gene: SGD:RSP5; PUB1; NPI1; MIPS:YER125w
A;Cross-references: MIPS:YER125w; SGD:S0000927
A;Map position: 5R
C;Function:
A;Description: involved in endocytosis of GAP1 protein and FUR4 protein; binds and ubi
C;Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C;Keywords: ligase
F;229-266/Domain: WW repeat homology <WW1>
F;331-368/Domain: WW repeat homology <WW2>
F;387-424/Domain: WW repeat homology <WW3>
Query Match 35.5%; Score 1377; DB 1; Length 809;
Best Local Similarity 37.0%; Pred. No. 4.8e-92;
Matches 313; Conservative 131; Mismatches 238; Indels 164; Gaps 24;
QY 3 SSIKILTVLCAKNIKKAKKDFRLPPFAKIVDGGSGCHSTDTVQNTLDPKWNHGYDL-Y 61
Db 3 SSISVXL--VAESLYKRDVFRSPDPFPAVLIDGY-QTKSTSAKKTLNPNYNETFKFDD 59
QY 62 VGKTDSTITISVNNHKKHKOGAGFLGCVRL-LSNAISRLKDT-----TGYOR-----LDL 110
Db 60 INENSILITIOVFDQKPK-KKKDQGLGVNVRVGVLDGLDEDTATSSGRPREETITRDL 118
QY 111 CKLNPSDTPAVRGQIVV-----SIQTRDRIGTGVSVVDCRGLL 148
Db 119 KKSND--DGMVSGRLIVLSKLPSSPSHQAQSGHTASSNTSTSTTNGHSTSS---T 173
QY 149 ENEGTVYEDSGPRLSCFMEEFAP-----YTDSTGAAA----- 182
Db 174 RNHSTSPSRGTAQAVESTLQSGTTAATNTATTSRSTNSTSSATROYSSPEDQYGRLLP 233


```

183 -----GGGNCRFYE-----SPQDORLOAQRINPDRVGRSLQTPQNR-----PHG 222
234 GWERRTDNFGRTYYVDHNRTRTTTWKRPDLQ--TEAERGQNLNANTELERRQHRGRTLPQG 292
223 HOS-----PELPEGEYQORTTVOGVYFLHT 247
293 SBDNSVTVQGGGNIIPVNGAAAAAFAATGGTSGLGELPSSGWEORFTPEGRAVFDH 352
248 QTCGVSTWHDPRIPRDLNSYN-----CDELGLPPGWESTVSGRIYFVDHNRRT 297
353 NRTTITWVDRPQQVIRTVPTNTTIQQOPVQLGELPSSGHEMLTNTARVYFVDHNRKT 412
298 TQPTDRLHHIWHQCQLKEPQPPPLPSEGSLEDEELPAQRYERDVLQKLVLRHELST 357
413 TTWDDPR-----LPS--SL--DQNVF--QYKCDPRKVIYFRSPAL 448
358 QPQAGHCRIEVSREIPEESYQIMKRPDKLRLMKVFRGEGLDYGGVAREMIYLL 417
449 -RLPGCHIKVRRKNIFEDAYQEIQRQTFEDLKKRLMIKFDGEGLDYGGVAREMIYLL 507
418 CHEMLNPFYGLFOYSTDNITMQLNPDSINPDHLSYFHFVGRINGLAHFGHYINGFT 477
508 SHEMFNPFYCLFEYSAYDNTYQINPNSGINPEHLNYPKFGVGVVGLGVFRRLDAPFFV 567
478 VPYKOLLGKPIQLSLESVDPELHKSILVILENDITPVLDHTFCVEHNAFRGRILOHELK 537
568 GAIYKMLKRVVLDQVEGDAEVNLSANWLENSIDGVLDITFSADDERFGEVTVDLK 627
538 PNGRNVPVTEENKEVRLVYVNRFRMGTEAQLQKGFNELIPQHLLKFPDQKELELI 597
628 PDGRNTEVTDGNEKEYVELYQWRIVDVQEQFKAFMDGFGNELIPEDLVTVFDESELELL 687
598 IGGLDKIDLDNKSNTRLKHCVADSNIVWFVQAVETFEERRARLLQFVGTGRVPLQ 657
688 IGGIAEIDEDKXHDYRGYQESDEVIQFWKCVSEWNEQEARLLQFTGTSTRIPVNG 747
658 FKALQGTGAGPRLFTIHLIDANTNLNPKAHTCFNRIDIPYESYKLYEKLTAVEET 717
748 FKDLQGSDD--GFRRTTIEKA--GEVQQLPKSHTCFNRVLDVQVYDYSMKQKLTAVEET 803
718 CGFAVE 723
804 IGFQGE 809

;RESULT 5
;Species: Schizosaccharomyces pombe
;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
;Accession: T39585
;Volckart, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
;Submitted to the EMBL Data Library, August 1997
;Reference number: 221865
;Accession: T39585
;Status: Preliminary; translated from GB/EMBL/DBJ
;Molecule type: DNA
;Residues: 1-786 <VOL>
;Cross-references: EMBL:299759; FIDN: CAB16903.1; GSPDB: GN00067; SPDB: SPBC16E9.11c
;Experimental source: strain 972h-; cosmid c16E9
;Genetics:
;Gene: SPDB: SPBC16E9.11c
;Map position: 2
;Introns: 60/2
;Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
;236-273/Domain: WW repeat homology <WWR1>
;306-343/Domain: WW repeat homology <WWR2>
;364-401/Domain: WW repeat homology <WWR3>

Query Match 34.5%; Score 1341; DB 2; Length 786;
Best Local Similarity 36.6%; Pred. No. 1.9e-89;
Matches 304; Conservative 128; Mismatches 236; Indels 162; Gaps 21;

```

```

QY 6 KIRLTVLCAKNLAKKOFERLPDPFAKIVVDGSGQCHSTDTVKNTLDPKKNOHYDLYVGKT 65
DB 7 RVRVYVAAGLAKSRDLFRQDFALITVDGB-QHTTKVKKSNPNYWGFEVTVKPS 65
QY 66 DSITISVNNHKKIHKQAGFLGCVRLLSNAISRLKDTGYOR-----LDLCK-LNPSDD- 119
DB 66 SVISIRLFDQKF-KKXQGLGLVGF-----RMEVGSFRSNREVSLRPLKKSSTN 118
QY 120 -AVGQIVW-----SL 129
DB 119 LSVLGNLVLKVAPSKIRAPAGNHSSTANTTSTPTTTTARTTRTPRTATNTSNQST 178
QY 130 QTRDRIGTGGVDCRGLLENEGTVYEDSGFRPL-----SCFMEBPA-----PYTOS 177
DB 179 SNSITNGTSAATNGTGTGAGTGASHRSS-----PVNRTQNTSALSNSNAHIMSFEQ 234
QY 178 TGAAGG-----GNCRFYE-----SPQDORLOAQRINPDRV 209
DB 235 YGRLLPPGWERRADSLGRYYVDHNRTRTTWTRPASSTNPNVHNTSSDSORLNQHRPLD- 293
QY 210 RGSLLQTPONRPHGHQSP---ELPEGEYQORTTVOGVYFLHTQGVSTWHDPRIP- 260
DB 294 -----DSNLSLQSDSGNDLPGWEMRYTDGRPYFVDHNRTRTTWVDPNPLVRPNG 346
QY 261 -----RLNSVNCDELGLPLPPGWESTVSGRIYFVDHNRTRTQPTDPELHHIMHQCQ 314
DB 347 GSSTVGSLMQPSLSLGLPLPSGHEMLTNSARVYFVDHNRTRTTWDDPR- 397
QY 315 LKPEQPLPLPSEGSLEDEELPAQRYERDVLQKLVLRHELSSLQPOAGHCRIEVSREI 374
DB 398 -----PSALDQDVP--QYKCDPRKVIYFRSPGM-RPLPGCNVKKVRCHI 441
QY 375 FEESYQIMKRPDKLRLMKVFRGEGLDYGGVAREMIYLLCHEMLNPFYGLFOYSTD 434
DB 442 FDSYAEIMRYSAHDLKRLMIRFDGEGLDYGGLSREFFLLSHKMFDFIYCLFEYSAV 501
QY 435 NYMLQINPDSSINPDHLSYFHFVGRINGLAHFGHYINGFTVPFYKOLLGKPIQLSDL 494
DB 502 DNYTQINPHSINPEHLNYPKFGVGVVGLGVFRRLDAPFFVYSLYKLLRKKVSLADM 561
QY 495 ESDVPELHKSILVILENDITPVLDHTFCVEHNAFRGRILOHELKPNGRNVPVTEENKEV 554
DB 562 ESIDAEFYSLKWLLENDITGLDITFSVEEDHFGVTRVELITNGENIEVTEENKKYV 621
QY 555 RLYVNRFRMGTEAQLQKGFNELIPQHLLKFPDQKELELIIGGLDKIDLDNWKGNTR 614
DB 622 DLVTEWRVSKRVQEQFNAPYSGFVELVSPDLVNVFDERLELLIGGISDVDEWKSHTE 681
QY 615 LKHCVADSNIVWFVQAVETFEERRARLLQFVGTGRVPLQGFALQGSSTGAGPRLFT 674
DB 682 YRYIATDPVKKWFWEIIAGWKVNEKSKLLQFATGTSRIPVNGFRDLQGSDD--GPRKFT 738
QY 675 IHLIDANT-DNLNPKAHTCFNRIDIPYESYKLYEKLTAVEETCGFAVE 723
DB 739 IE--KAGTPDQFVAHTCFNRVLDVQVYDYSMKQKLTAVEETVGFQGE 786

RESULT 6
T46412
ubiquitin-protein ligase (EC 6.3.2.19) NEDD4 - human (fragment)
N;Alternate names: hypothetical protein DKFZp334P2422.1
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 03-Jun-2002
C;Accession: T46412
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: 223034
A;Accession: T46412
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-820 <AAA>
A;Cross-references: EMBL:AL137469

```

Experimental source: adult testis; clone DKFzp34P2422

Genetics:

Note: DKFzp34P2422.1

Gene: GDB:NEDD4

Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin

Keywords: ligase

142-179/Domain: WW repeat homology <WR1>

342-379/Domain: WW repeat homology <WR2>

393-430/Domain: WW repeat homology <WR3>

489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 32.0%; Score 1241.5; DB 2; Length 820;

Best Local Similarity 34.9%; Pred. No. 3.7e-82;

Matches 294; Conservative 116; Mismatches 230; Indels 203; Gaps 21;

43 TDVNTLDPKQNHVDLYVGKTD-SITSVWNKKHKQAGFLGCVRLLSNAI----- 97

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 97

12 TKTIKTLPKQNEEFVFNPSNHLLEFEVDENLRTRD---FLGQVDVPLSHLPTD 68

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 68

98 -----SRLDGTYGRLDCKLNPSDTPDAVGQIVSLQTRDRTG 137

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 137

69 PTWERPWFQDFLLRPRSHKSRVK--GFRLKWAYNPKNG-----GQDENSDQDDMEH 121

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121

138 GGSVYDCRG-----LLENQGVY----- 155

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 155

122 GWEVVDNDSASQHEBELPPPLPGWEKVDNLGRTYVYVNNRRTTQWHRPDLMDVSE 181

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181

156 -----EISGP-----GRPLSCFWEPPAPYTDSTGAA-- 181

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181

182 SDNNIRQINQEAHRRFRSRHISELEPEPSGGVDPPEWETISEVNIAGDSLGLALP 241

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 241

182 ---AGGNCRFVSPSQD-----ORLQARLN----- 206

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 206

242 PPPASPSRSTSPQELSELSRLRLQITPDSNGEFSLLIQEPSSRLRSCSVTDAVEQGH 301

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 301

207 -----PDVRSLSQTPQNRPHQSPH-----LPEYGEQRTTVQGVYEL 245

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 245

302 LPPGAKDSPVRAVKDTLSLPQSPQSPYNSPKQHKVTQSLFPPGWEWRIAPNGRPFFI 361

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 361

246 HTQGVSTWHDPRIP-----RLNSVNCDELGLPPLPGWEVSTVSGRIYFVDHNRRTQF 300

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300

362 DHNTKTTTWEDPLRLKFFVHNRKSTSLNPDLGLPLPGWEERILHLDGRFTYDHSNKITQW 421

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 421

301 TDPLRLHNMHCQQLKEPSQPLPLPSGSLDEBELPAQRYERDLVKLKLRLHLSLQOP 360

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360

422 EDPRLN-----PAITG-----PAVPYSREFKQKYDFRKLKPKPAD 458

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 458

361 QAGHCRIVRSREEIFERSYRQIMKX-RPKDLKKELMVKFRGEGLDYGGVAREWLYLLCH 419

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 419

459 IPNRFEMKLRHNNIFESYRIRMSVKRDPVVKARLWIEFESEKGLDYGGVAREWFFLSK 518

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 518

420 EMLNPYGLFOYSTDNIMQLNPDSI-RPDHLSYFHFVGRIMGLAVFHGHVINGGFTV 478

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 478

519 EMFNPNYGLFEYSATDNYTLQINPNSGLCNEHLSYFTFGRVAGLAVFHGKLLDGFTR 578

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 578

479 PRYKOLLGKPLQLDSVDPDLKSLVILLENDITVPLDHTFCVSEHNAFGRILQHELAP 538

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 538

579 PRYKMLGKQLTLNDSVDEYNSLKNLENDPTB-LDLMCIDBEENFGQYQVDLAP 637

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 637

539 NGRNPVPTENKKEYRLVYNNRMFRGIEAQFLAQGFNELIPQHLKLPDQKELELII 598

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 598

638 NGSEIMVTNENKREYIDLVIQWRFNVRVQKQNAFLEGFTPELLIDLIKIFDENELLELM 697

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 697

599 GLDKIDLNDKSNTRLKHCVADSN-IVRPMQAVETFDERRARLLQFTVGSSTRYPLOG 657

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 657

698 CGLGDVDVNDNRQHSIYKNGYCPNHPVQWFWKAVLLMDAEKIRLLQFVIGTSRVPNG 757

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 757

658 FKALQGSTGAAGPRFLTTHLIDANTDLNPKAHTCFNRIDIPPVESYEKLEKLLTAAEET 717

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 717

758 FAELYGSN---GPQLFTTEQW-GSPEKLPRAHTCFNRLDLPPEYTFEDREKLLMAVENA 813

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 813

718 CGF 720

Y |||

Db 814 QGF 816

RESULT 7

S70642

ubiquitin ligase Nedd4 - rat (fragment)

C/Species: Rattus norvegicus (Norway rat)

C/Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 18-Aug-2000

C/Accession: S70642

R/Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.

EMBO J. 15, 2371-2380, 1996

A/Title: WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na(+

A/Reference number: S70642; MUID:96221297; PMID:8665844

A/Accession: S70642

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-887 <STA>

A/Cross-references: EMBL:U50842; NID:g1293646; PIDN:AAB48949.1; PID:g1293647

C/Genetics:

A/Gene: Nedd4

C/Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubi

F/54-157/Domain: protein kinase C C2 region homology <XC2>

F/246-283/Domain: WW repeat homology <WW1>

F/402-439/Domain: WW repeat homology <WW2>

F/459-496/Domain: WW repeat homology <WW3>

F/555-881/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 31.6%; Score 1227.5; DB 2; Length 887;

Best Local Similarity 33.9%; Pred. No. 4.4e-81;

Matches 296; Conservative 127; Mismatches 224; Indels 227; Gaps 28;

QY 7 IRLTVLCARNLAKKQFFRLPDPFAKIV---VDGSGQCHSTDTVKNILDPKN----- 55

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 55

Db 77 VRVYKTIAGILAKKQILGASDPYVRVTLVDPMSGVLTSVQTKIKSLNPKWNEELFRV 136

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 136

QY 56 ---QH---YDLVVGKTD---SITISVWN-----HKIKHKQG 83

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 83

Db 137 LPQHRILFEVDENLRTRDFFLGQVDVPLPLPTENRMERPYTFKDFVLRPSHKSrv 196

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 196

QY 84 AGFLGCVRLLSNAISRLKDT----- 103

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 103

Db 197 KGYL-----RLKMTYLPKNGSDDENADQAELEPGWVLLDQPDAAATHLQHPPEP 245

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 245

QY 104 -----GY-QRLDLC-----KLNPSD--TDAVGOIVV---SLQTRDRI 135

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 135

Db 246 SPLPPGWEERQDVIGRTYVYVNNHESRTQWRKSPEDDITDDENGDIQLQAHGAFTRRQI 305

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 305

QY 136 GTGGSVVDCRG-----LLENQGVYEDSGGRPLS-----C 166

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 166

Db 306 SE-----DVDGPDNHESPENWEIVREDENTIIYSGAVQSPSPSGHPDVQVRLAEELDTLTL 360

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360

QY 167 FMEEPA---PYTDSTGAAAGGN--CRFVESPSQORLQAOQLRNPVYRGSILQTPQNRPH 221

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 221

Db 361 MYGNPATSQVTSNHSRSGSSQTCIFEEQPTLVLLFT----- 400

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 400

QY 222 GHOSPELPEGEORTTVQGVYFHLTQTGVSTWHPD-----RIPRDLNS---VNCD 270

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 270

Db 401 ---SSGLPFGWEKQDGRGSIYVDHNSKTTTWSKPTMQDDPRSKIPAHLRGKTPVDSND 457

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 457

QY 271 LGELPPGWEVRSTVSGRIYFVDHNRRTQFTPRLHHNMHCQQLKEPSQPLPLPSEGLS 330

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 330

Db 458 LGELPPGWEERTTDRGVFFNNHNIKKQTQWEDPRQNV-----AITGPAEP----- 503

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 503

QY 331 EDELPACRYERDLVKLKLRLHLSLQOPQAGHCRIVRSREEIFERSYRQIMKVRPKD- 389

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 389

Db 504 -----YSRIYKRYEPPFRKLKKQTDIPNKKEMKLRANILEDSYRRIMGVKRADF 554

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 554

QY 390 LKGLMYKVFGESEGLDYGGVAREWLYLLCHEMLNPNYGLFQYS-TDNIYMLQINPDSST- 447

615 NEDHLSYFKFGRVAGMAYVHGKLLDGFIRPFYQOMLQKLIITHDMESVDSEYYSRLW 674
508 ILENDITPVLDTFCVEHNAFGRILQHELKPNRNPVPTTEENKEYVRLVYNWRFMRGIE 567
675 ILENDPTE-LDLRFIIDEELFQTHQELKGTGSSVVVTKNKKKEYLYLVQWRFVNRIO 733
568 AQFLAQKGFNELIQLHLLKPDQKLELILIGGLDKIDLNDKSNTRKHCVA-DSNIVR 626
734 KQWAAPEKGFELIPQDLIKI FDEBELLMCGLDVNDWRREHTKYNGYSUNHQVH 793
627 WFOVAETEDEERRARLLQVFGTSRVLPQGFKALQSGTGAAGPRLEFTHLIDANTONLP 686
794 WFKAVLWMDSEKIRLLQVFGTSRVPMNGFAELYGSN---GQSFTEQW-GTPDKLP 849
687 KAHTCNRRIDIPYESYEKLYKLTAVETGCF 720
850 RAHTCFNRLLDPYESFDELWDLQWAIENIQGF 883

SULT 8
3196
DD-4 ORF - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Aug-2000
Accession: I83196
Kumar, S.; Tomooka, Y.; Noda, M.
ochem. Biophys. Res. Commun. 185, 1155-1161, 1992
Title: Identification of a set of genes with developmentally down-regulated expression
Reference number: I80167; MUID:92328780; PMID:1378265
Accession: I83196
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: mRNA
Residues: 1-708 <RES>
Cross-references: GB:D10714; NID:G220508; PID:G220509
Genetics:
Gene: NEBD-4
Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
40-77/Domain: WW repeat homology <WW1>
196-233/Domain: WW repeat homology <WW2>
251-288/Domain: WW repeat homology <WW3>
347-662/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 27.1%; Score 1051.5; DB 2; Length 708;
Best Local Similarity 36.1%; Pred. No. 2.1e-68;
Matches 244; Conservative 102; Mismatches 194; Indels 135; Gaps 19;
50 LDPKWNQHYDLYVGKTDSTITISVNNHKKIHKQAGFLGCVRLLSNAISRLKXDTGYRLD 109
42 LPPGWEERQDV-LGRYYV-----NHES-----RRTQWKR-- 70
110 LCKLNPS-----DTDAVGGQIVSLQTRDRIGTGGSVDCRGLLEN-----EGTV 154
71 -----PSPDDDLTDEDNDMDQLQRAFTTRQISBDVGPDRNSPFWKWEIVREDENTE 125
155 YEDSGPRGLSCFME-----BPA---PYTSTGAAAGG--NCRFVES 192
126 YSGAVQSPGSHGIDVQTHLAEBENTELAVCGNPATSPQVTSNHSRSGSLQTCIFEEQ 185
193 PSQORLOAQRLRNPDRVRSGLQTPNRPNGHQSPPELPEGVEQRTTVQGVFLHTQTGVS 252
186 PTLFVLLPT-----SSGLPPGWEERQDGRSYYVDHNSKTT 222
253 TWHDP-----RIPRDL-NSVNCDELGLPPGWEVRSTVSGRIYFVDHNNRTQFTDP 303
223 TWSKETMODDPRSKIPAHLRKTDNSDLGLPFCWEERTHTDGRVFFINHNKKTOWEDP 282
304 RLHHTMHQCLKPSQPLPLPSGSGLEDEBELPAQRVERDLVKLKVLRHLSLQOPQAG 363
283 RLQNV-----AITGPAV-----YSRDYKRYEYFFRKLKQTDIPN 319
364 HCRIVSREEIFEESYROIWKRPKD-LKKRLMYKFRGESGLDYGGVAREWLYLLCHEML 422
320 KPMEKLRANILDESYYRIMGVKRADLLKARLWIEFPGKGLDYGGVAREWFLISKEMF 379

423 NPYGLTFOYSTDNIMQLINPDSSI-NPDHLSYFHFVGRIMGLAVFHGHIYINGGFTVPY 481
380 NPYGLTFOYSTDNIMQLINPDSSI-NPDHLSYFHFVGRIMGLAVFHGHIYINGGFTVPY 439
482 KQLLQKPIQISDLESVDPELHSLVWILENDITPVLDTFCVEHNAFGRILQHELKPNR 541
440 KMWLQKLTITLHDMESVDSEYYSRLWILENDPTE-LDLRFIIDEELFQTHQELKGTGS 498
542 NPYTEENKKEYVRLVYNWRFMRGIEAQFLAQKGFNELIQLHLLKPDQKLELILIGGL 601
499 EIVTVNKKKEYLYLVQWRFVNRIOKMAAFKEGFELIPQDLIKI FDEBELLMCGSL 558
602 DKIDLNDKSNTRKHCVA-DSNIVRWFQAVETFDERRARLLQVFGTSRVLPQGFKA 660
559 GDVDVNDWRREHTKYNGYSUNHQVHFWKAVWMDSEKIRLLQVFGTSRVPMNGFAE 618
661 LOGSTGAAGPRLEFTH 675
619 LYGSN---GQSFTEV 630

RESULT 9
T37900
probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C;Accession: T37900
R;Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21752
A;Accession: T37900
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-671 <RES>
A;Cross-references: EMBL:AL117390; PIDN:CAB55856.1; GSPDB:GN00066; SPDB:SPAC1805.15C
A;Experimental source: strain 972h-; cosmid c1805
C;Genetics:
A;Gene: SPDB:SPAC1805.15C
A;Map position: 1
A;Intons: 60/2; 105/1; 639/2
C;Superfamily: WW repeat homology
F;242-279/Domain: WW repeat homology <WWR>

Query Match 26.4%; Score 1026.5; DB 2; Length 671;
Best Local Similarity 32.9%; Pred. No. 1.3e-66;
Matches 248; Conservative 128; Mismatches 253; Indels 125; Gaps 18;
6 KIRLTVLCAKRLAKDFFRLPDPFAKIVVDGSGGCHSTDTVKNITLDPKWNQHYDLYVGKT 65
7 EVQLTILHVEGLWKNGLRLSLKPYLLISVD-DDQFIKTNVASGTLRLSWGFTQKLTVSPQ 65
66 DSTTISVNNHKKIHKQAGFLGCVRLLSNA---ISRLKDTGYORLDLCKLNPSDTPDAVR 122
66 SILLGLQFDEKQ-KNETSDGFLGAAVNVNFPNFKDDYKTRITL-----RSPSGSYR 120
123 GQIVWSLQTRDRIGTGGSVVDCRGLLENEGTVYEDSGPRPLSCFMEEPAP-----YTD 176
121 GSVVCLFKR-----SKFLPEELPADKSOICTD 147
177 STGAAG-----GGNCRFVESP---SQDQRLQAOQLRNPDRVRSGLQTPNRPNGHQVS 225
148 IIDDASGCWETRIDFEGHYVYKSPQSVISAISHEKLEN-----LTFK----- 192
226 PELPEGVEQRTTVQGVFLHTQTGVSWM-----HDPRIPLDINSVNCDEL 271
193 -QLKEVFSQ-----FLFNQSKSLKINLEYKVIKHLLEHYPLALSVRQOVAVEK 241
272 GPLPPGWEVRSTVSGRIYFVDHNNRTQFTDRLHHIMHQCQLKEPSQPLPLPSGSGLE 331
242 GPLPAGWEMELSEDYHYFVDHSTKTITWSDPRDNV-----ASDSVS 284
332 DEELPAQRVERDLVKLKVLRHLSLQOPQAGHCRIVSREEIFEESYROIWKRPKD 389

RESULT 12
H96599
protein F14J16.10 [imported] - Arabidopsis thaliana
C1Species: Arabidopsis thaliana (mouse-ear cress)
C1Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C1Accession: H96599
R1Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cresasy, T.H.; De

sen, N.F.; Hughes, B.; Huizar, L.
ture 408, 816-820, 2000
Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
zo, M.; Rooney, T.; Rowley, D.; Sakano, H.
Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
r, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
Reference number: A86141; MUID:21016719; PMID:11130712
Accession: H98599
Status: preliminary
Molecule type: DNA
Residues: 1-4056 <STO>
Cross-references: GB:AE005173; NID:g8778329; PIDN:AAF79338.1; GSPDB:GN00141
Gene: F14J16.10
Map position: 1
Query Match 20.2%; Score 784; DB 2; Length 4056;
Best Local Similarity 39.4%; Pred. No. 9.5e-48;
Matches 172; Conservative 65; Mismatches 138; Indels 62; Gaps 7;
343 DLVOKLVLRHLSLQOPO--AGHCRIEVSREIEPESYROIIMKVRPKDKRLMKVFRG 400
3624 DFNKKAYFRSRIRHQDQHSGLPLISVRRAYVLEDSYNQLRMRSPODLKGRNLNVQFQG 3683
401 EEGLDYGGVAREWLYLLCHEMLNPYYGLFOYSTDNIMQLNPSSINPDHLSYHFVGR 460
3684 EEGIDAGLTREWYQLLSRVIFDKGALLFT-TVGNDATFQPNPNSVYQTEHLSYFKEVGR 3742
461 IMGLAVPHGHVINGGTFVPFYKOLGKPIQLSDLESVDPELHSLVWLENDITPVLDDHT 520
3743 MVAKALFDGQLDVFYTRSFYKHILGVKTYHDIKAVDPDYKYNLWLENDVSDILDLT 3802
521 F-----CVE---HNAFGRILQHELPKNGRN 542
3803 FMSDADEKHILYEKTEVTYDYLKPGGRNIRVTEETKHEYVDLVADHILTSAIRPQINAF 945
543 VPVTEENKKEYLYNWRNFMRGIEAQFALOKGNFELIPQHLKLPFDQKLELIIGLD 602
3863 IRVTEETKHEYVDLVAGHILTNAIRPQINAFLEGNELIPRELVSIFNPKLELLISGLP 3922
603 KID-----LNDKSNRLKHCVADSNIRVFWQAVETDEERRARLLQFV 647
3923 EIDCKLDSIOYLCAVPLIDDLKANTYTSYTAGSPVIHFWVWKAFSKEDMARFLQV 3982
648 TGSTRVPLQFGKALQSGTGAAGPRFLTIHLIDANTNLPAKTCNFRIDIPPYSEYKLY 707
3983 TGTSKVPLEGFKALQ---GISGQRLQHKAYGAPERLPSAHTCFNQLDLPYQSKQLQ 4039
708 EKLTAVETC---GFA 721
4040 ERLLLAHEASGEGGFA 4056
RESULT 13
31491
ubiquitin-protein ligase homolog F1707.15 - Arabidopsis thaliana
Species: Arabidopsis thaliana (mouse-ear cress)
Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 04-Mar-2000
Accession: T01491
Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li,
z, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N
ubmitted to the EMBL Data Library, June 1998
Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.
Reference number: Z14334
Accession: T01491 from GB/EMBL/DBJ
Status: translated
Molecule type: DNA
Residues: 1-1126 <VYS>
Cross-references: EMBL:AC003671; NID:g2833627; PID:g3176690; GSPDB:GN00059; ATPSF:F1707
Experimental source: cultivar Columbia
Genetics:

A:Gene: ATPSF:F1707.15
A:Map position: 1
A:Introns: 118/3; 224/3; 292/1; 373/3; 457/2; 826/3; 875/3; 902/3; 1057/3; 1093/2
C:Superfamily: ubiquitin-protein ligase homology
F:756-1120/Domain: ubiquitin-protein ligase homology <UBI>
Query Match 20.1%; Score 780.5; DB 2; Length 1126;
Best Local Similarity 40.1%; Pred. No. 2.4e-48;
Matches 170; Conservative 67; Mismatches 138; Indels 49; Gaps 6;
343 DLVOKLVLRHLSLQOPO--AGHCRIEVSREIEPESYROIIMKVRPKDKRLMKVFRG 400
707 DFNKKAYFRSRIRHQDQHSGLPLISVRRAYVLEDSYNQLRMRSPODLKGRNLNVQFQG 766
401 EEGLDYGGVAREWLYLLCHEMLNPYYGLFOYSTDNIMQLNPSSINPDHLSYHFVGR 460
767 EEGIDAGLTREWYQLLSRVIFDKGALLFT-TVGNDATFQPNPNSVYQTEHLSYFKEVGR 825
461 IMGLAVPHGHVINGGTFVPFYKOLGKPIQLSDLESVDPELHSLVWLENDITPVLDDHT 520
826 MVAKALFDGQLDVFYTRSFYKHILGVKTYHDIKAVDPDYKYNLWLENDVSDILDLT 885
521 FCEVHNAPGRIL-----QHELPKNGRNVPVTEENKKEYVRLYVNRNFMRGIEAQFLAL 573
886 FMSDADEKHILYEKTEVTYDYLKPGGRNIRVTEETKHEYVDLVADHILTSAIRPQINAF 945
574 QKGFNELIPQHLKLPFDQKLELIIGLDKID----- 605
946 EELNELIPRELVSIFNPKLELLISGLPELDCKLAPYFIFLHFAYSFKVITILLSV 1005
606 -----LNDKSNRLKHCVADSNIRVFWQAVETDEERRARLLQFVTSRVPLOGFKA 660
1006 FFCFLVDLKKANTYTSYTAGSPVIHFWVWKAFSKEDMARFLQFVTGTSKVPLEGFK 1065
661 LGSTGAAGPRFLTIHLIDANTNLPAKTCNFRIDIPPYSEYKLYEKLITAVEET--- 717
1066 LQ---GISGQRLQHKAYGAPERLPSAHTCFNQLDLPYQSKQEQVQERLLLAHEANEG 1122
718 CGFA 721
1123 FGFA 1126
RESULT 14
S69625
hypotheical protein YDR457w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C:Accession: S69625
R; Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A:Reference number: S69554
A:Accession: S69625
A:Molecule type: DNA
A:Residues: 1-3268 <DIE>
A:Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64910.1; PID:g927738; MIPS:YDR45
C:Genetics:
A:Gene: SGD:TOM1
A:Cross-references: SGD:S0002865; MIPS:YDR457w
A:Map position: 4R
Query Match 19.8%; Score 769; DB 2; Length 3268;
Best Local Similarity 43.3%; Pred. No. 8.4e-47;
Matches 152; Conservative 70; Mismatches 125; Indels 4; Gaps 2;
367 IEVSREIEPESYROIIMKVRPKDLK-KELIMVKFRGEGLDYGGVAREWLYLLCHEMLNPY 425
2913 ITVRRQVFLDSYRALFPKTNDEIKNSKLEITFKGESGVADAGVTREWYQVLSRQMENPD 2972
426 YGLFOYSTDNIMQLNPSSINPDHLSYHFVGRIMGLAVPHGHVINGGTFVPFYKQLL 485
2973 YALFLVPSDKTTFHPNRTSGINPEHLSFFKPIGMIIGKAIKRDQCFLDCHFSREYKNI 3032

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

n on: February 20, 2004, 15:25:26 : Search time 10.8131 Seconds
(without alignments)
3144.377 Million cell updates/sec

tle: US-10-009-945-2

efect score: 3884

quence: 1 GGSSIKIRLTVCANLAKK.....EKLYEKLITAVEETCGFAVE 723

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 127863 seqs, 47026705 residues

tal number of hits satisfying chosen parameters: 127863

nimum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

tabase : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match	Length	DB ID	Description
1	3855	99.3	757	1	SUFL1_HUMAN
2	3548	91.3	731	1	SUFL1_XENLA
3	3249.5	83.7	619	1	SUFL1_MOUSE
4	3001	77.3	748	1	SUFL2_HUMAN
5	1492.5	38.4	767	1	PUB1_SCHPO
6	1377	35.5	809	1	RSP5_YEAST
7	1259	32.4	927	1	NED4_HUMAN
8	1225	31.5	957	1	NED4_MOUSE
9	627.5	16.2	310	1	URB1_RAT
10	534	13.7	875	1	UE3A_HUMAN
11	483	12.4	885	1	UE3A_MOUSE
12	481.5	12.4	892	1	HUL4_YEAST
13	449.5	11.6	1050	1	HER3_HUMAN
14	358	9.2	1992	1	TRIB_HUMAN
15	354	9.1	910	1	HUL5_YEAST
16	278.5	7.2	2799	1	EDD_HUMAN
17	277	7.1	920	1	EDD_RAT
18	262.5	6.8	1483	1	UFDA_YEAST
19	257	6.6	1647	1	YDE1_SCHPO
20	238.5	6.1	472	1	YAPI_MOUSE
21	238.5	6.1	1620	1	HED1_HUMAN
22	222	5.7	2895	1	HYD_DROME
23	162	4.2	448	1	YAPI_CHICK
24	161.5	4.2	454	1	YAPI_HUMAN
25	128	3.3	1813	1	UN13_CAEEL
26	119	3.1	575	1	BAG3_HUMAN
27	116.5	3.0	577	1	BAG3_MOUSE
28	114.5	2.9	826	1	RSBG_HUMAN
29	114	2.9	593	1	CNE5_HUMAN
30	113	2.9	1021	1	YPT7_CAEEL
31	111.5	2.9	672	1	KPCA_HUMAN
32	111.5	2.9	672	1	KPCA_RABIT
33	111.5	2.9	672	1	KPCA_RAT

34	111.5	2.9	5147	1	PCLO_HUMAN
35	110.5	2.8	597	1	IXR1_YEAST
36	110	2.8	658	1	KPC1_LYPTI
37	109.5	2.8	537	1	CNE3_HUMAN
38	109	2.8	694	1	RP3A_HUMAN
39	109	2.8	704	1	RP3A_BOVIN
40	109	2.8	915	1	PDB2_ARATH
41	108	2.8	506	1	PYCA_ARCFU
42	108	2.8	665	1	RNR2_LACLA
43	108	2.8	1013	1	PTPX_MACNE
44	108	2.8	1178	1	PXC_MOUSE
45	108	2.8	5120	1	PCLO_CHICK

ALIGNMENTS

RESULT 1
SUFL1_HUMAN STANDARD; PRT; 757 AA.
AC Q9HCE7; O75853; Q9UUT8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
DE protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
DE (hSMURF1).
GN SMURF1 OR KIAA1625.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=93385348; PubMed=10458166;
RX Zhu H., Kaysak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
RT pattern formation.";
RL Nature 400:687-693 (1999).
RN [2]
RP SEQUENCE OF 20-731 FROM N.A. (ISOFORM SHORT).
RA Stoneking T., Bauer C., O'Neal D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hiroseawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res 7:273-281 (2000).
CC -!- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q9HCE7-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q9HCE7-2; Sequence=VSP_006812;
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).


```
-----
EMBL; AF199364; AAF08298.2; -
EMBL; AC004893; AAC62434.1; -
EMBL; AB046845; BAB13451.1; ALT_INIT.
HSSP; Q13526; 1PIN.
MIM; 605568; -
GO; GO:0005622; C:intracellular; TAS.
GO; GO:0000211; P:protein degradation tagging activity; IDA.
GO; GO:0004842; P:ubiquitin-protein ligase activity; IDA.
GO; GO:00030154; P:cell differentiation; TAS.
GO; GO:0007398; P:ectoderm development; IDA.
GO; GO:0030514; P:negative regulation of BMP signaling pathway; TAS.
GO; GO:0006464; P:protein modification; TAS.
GO; GO:0006512; P:ubiquitin cycle; IDA.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 2.
SMART; SM00239; C2; 1.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 2.
PROSITE; PS00459; C2_DOMAIN 1; 1.
PROSITE; PS00004; C2_DOMAIN 2; 1.
PROSITE; PS0237; HECT; 1.
PROSITE; PS01159; WW_DOMAIN 1; 1.
PROSITE; PS0020; WW_DOMAIN 2; 2.
UBI conjugation pathway; Ligase; Repeat; Alternative splicing.
DOMAIN 1 99 C2 DOMAIN.
DOMAIN 234 267 WW 1.
DOMAIN 306 339 WW 2.
DOMAIN 420 757 HECT.
BINDING 725 725 UBIQUITIN.
VARSPPLIC 269 294 Missing (in isoform Short).
MUTAGEN 725 725 C->A; LOSS OF UBIQUITINATION CAPACITY.
SEQUENCE 757 AA; 86113 MW; 89A171CFC47B40E9 CRC64;
Query Match 99.3%; Score 3855; DB 1; Length 757;
Best Local Similarity 96.5%; Pred. No. 8.8e-272;
Matches 722; Conservative 0; Mismatches 0; Indels 26; Gaps 1;
2 GSSIKRLTVLCANLAKKDFRLPPAKIVVDGSGGCHSDTDTVNTLDPKWNQHYDLY 61
10 GSSIKRLTVLCANLAKKDFRLPPAKIVVDGSGGCHSDTDTVNTLDPKWNQHYDLY 69
62 VGKTDSTTISVWNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQLDLCKLNPSTDAV 121
70 VGKTDSTTISVWNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQLDLCKLNPSTDAV 129
122 RQIVVSLQTRDRICTGGSVDCRGLLENGEYVDSGGRLSCFMEEPAPYTDSTGAA 181
130 RQIVVSLQTRDRICTGGSVDCRGLLENGEYVDSGGRLSCFMEEPAPYTDSTGAA 189
182 AGGNCRCFVESPQDORLQAQLRNPEDVRSGLTQPNRPHGQSPQLPEGEYEQRTTVQGG 241
190 AGGNCRCFVESPQDORLQAQLRNPEDVRSGLTQPNRPHGQSPQLPEGEYEQRTTVQGG 249
242 VYFLHTGTGTVTHDPRI-----PDLNSVNCDELGLP 275
250 VYFLHTGTGTVTHDPRI-----PDLNSVNCDELGLP 309
276 PGWEVRVTSGRIVYVDHNNRTTQFTDPLHIMNHQCOLKEPSPQLPSPGSGLEDEEL 335
310 PGWEVRVTSGRIVYVDHNNRTTQFTDPLHIMNHQCOLKEPSPQLPSPGSGLEDEEL 369
336 PAQRYEDLVQKLVRLHELSLQOPAGHCRIVSREBIFEESYQIMKRPKDKGLM 395
370 PAQRYEDLVQKLVRLHELSLQOPAGHCRIVSREBIFEESYQIMKRPKDKGLM 429
396 VKRGEGLDYGGVAREWLYLLCHEMLNPYGLFOYSTDNLYMLQINPDSSINPDHLSYF 455
-----
```

```
Db 430 VKRGEGLDYGGVAREWLYLLCHEMLNPYGLFOYSTDNLYMLQINPDSSINPDHLSYF 489
QY 456 HFVGRINGLAVFHGHINGGFTVPFYKQLLKPQLQSLDLESVDPELHKSILWILENDITP 515
Db 490 HFVGRINGLAVFHGHINGGFTVPFYKQLLKPQLQSLDLESVDPELHKSILWILENDITP 549
QY 516 VLDHTFCVEHNAFGRILOHELKPNGRNVPVTENKKEYVRLYVNRFRMEGIEBAQFLALQK 575
Db 550 VLDHTFCVEHNAFGRILOHELKPNGRNVPVTENKKEYVRLYVNRFRMEGIEBAQFLALQK 609
QY 576 GFNELIPOHLKPPDQKLELLIIGGLDKIDLNDWKSNTLKHCVADSNIVRWQAVETP 535
Db 610 GFNELIPOHLKPPDQKLELLIIGGLDKIDLNDWKSNTLKHCVADSNIVRWQAVETP 669
QY 636 DEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLKPAHTCFNRI 695
Db 670 DEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLKPAHTCFNRI 729
QY 696 DIPPYSEYKLYEKLTAVEETCGFAVE 723
Db 730 DIPPYSEYKLYEKLTAVEETCGFAVE 757

RESULT 2
SUFL_XENLA
ID SUFL_XENLA STANDARD; PRT; 731 AA.
AC O9PUN2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Smad ubiquitin ligase regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase) (xSMURF1).
DE SMURF1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RX [1]
SEQUENCE FROM N.A.
TISSUE=Blasutula;
MEDLINE=99385348; PubMed=10458166;
RA Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomson G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
pattern formation."
RL Nature 400:687-693(1999).
CC -!- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION. MAY
REGULATE ECTODERMAL DIFFERENTIATION AND PATTERN BY MODULATING BMP
SIGNALING AND MAY ENHANCE CELLULAR RESPONSIVENESS TO THE SMAD2
(ACTIVIN/TGF-BETA) PATHWAY.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM THE EGG STAGE TO THE SWIMMING
TADPOLE, WITH MAXIMUM LEVELS OBSERVED IN THE STAGES FROM EGG TO
GASTRULA. AT GASTRULATION DISTRIBUTED UNIFORMLY IN EMBRYONIC
ECTODERM AND INVOLUTING MESODERM, AND EXPRESSION GRADUALLY
LOCALIZES TO THE NERVOUS SYSTEM. AT EARLY TADPOLE STAGES EXPRESSED
IN THE CNS, EYE, BRANCHIAL ARCHES, KIDNEY AND SOMITES.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF169310; RAD52564.1; -
GO; GO:0005622; C:intracellular; TAS.
DR
```

```
GO; GO:0000211; F:protein degradation tagging activity; IDA.
GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
GO; GO:0030154; P:cell differentiation; IDA.
GO; GO:0007398; P:ectoderm development; TAS.
GO; GO:0030514; P:negative regulation of BMP signaling pathway; TAS.
GO; GO:0006464; P:protein modification; TAS.
GO; GO:0006512; P:ubiquitin cycle; IDA.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 2.
SMART; SM00239; C2; 1.
SMART; SM00119; HECT; 1.
SMART; SM00456; WW; 2.
PROSITE; PS00499; C2_DOMAIN_1; 1.
PROSITE; PS00004; C2_DOMAIN_2; 1.
PROSITE; PS02037; HECT; 1.
PROSITE; PS01159; WW_DOMAIN_1; 1.
PROSITE; PS00200; WW_DOMAIN_2; 2.
UBI conjugation pathway; Ligase; Repeat.
DOMAIN 1
99 C2 DOMAIN.
DOMAIN 233
266 WW 1.
DOMAIN 279
312 WW 2.
DOMAIN 394
731 HECT.
SEQUENCE 731 AA; 3CE98E512A42CE2C CRC64;

Query March 91.38; Score 3548; DB 1; Length 731;
Best Local Similarity 91.08; Pred. No. 1.5e-249;
Matches 659; Conservative 34; Mismatches 29; Indels 2; Gaps 2;

1 GGSIKIRLTVLCAKLNIAKDFRLLPDPFAKIVVDGSGGCHSTDTVKNLDPKMQHYDL 60
9 GGSIKIRLTVLCAKLNIAKDFRLLPDPFAKIVVDGSGGCHSTDTVKNLDPKMQHYDL 68
61 YVGKTDITTSVNHKKIHKQAGFLGCVRLLSNAISRLKDTGYORLDCIKLNSDTDA 120
69 YVGKVDITTSVNHKKIHKQAGFLGCVRLLSNAISRLKDTGYORLDCIKLNSDTDA 128
121 VVGIVVSLQTRDRIHTGGVDCRGLLENFTVYDSDGPRPLSCFMEEPAPYTDSTGA 180
129 VVGIVVSLQTRDRIHTGGVDCRGLLENFTVYDSDGPRPLSCFMEEPAPYTDSTGA 187
181 AAGGNCRFVESPQDQRLQALRNPDVRSGLQTPNPHQSPHSPGVEQRTTVQG 240
188 AAGGNCRFVESPQDQRLQALRNPDVRSGLQTPNPHQSPHSPGVEQRTTVQG 247
241 QVYFLHTQTVSTVHDPRIPLDNLNVNCDLGLPLPGWEVRSVTSVGRVYFVDHNNRTTQF 300
248 QVYFLHTQTVSTVHDPRIPLDNLNVNCDLGLPLPGWEVRSVTSVGRVYFVDHNNRTTQF 307
301 TDPRLHIMHQCCKPSQPLPLPSSGSLD-EELPAQRYERDLVQKLVLRLHLSLQ 359
308 TDPRLHIMHQCCKPSQPLPLPSSGSLD-EELPAQRYERDLVQKLVLRLHLSLQ 367
360 POAGHCRLEVSREIPEESVQIMKMPKDKRLVMKRGEGGLDYGVAWEWYLLCH 419
368 POAGHCRLEVSREIPEESVQIMKMPKDKRLVMKRGEGGLDYGVAWEWYLLCH 427
420 EMLNPYGLFYQSTNTYMLQINPDSSINPDHLSYHFVGRIMGLAVFHHYNGFTVP 479
428 EMLNPYGLFYQSTNTYMLQINPDSSINPDHLSYHFVGRIMGLAVFHHYNGFTVP 487
480 FYKQLGKPIQLSDLESVDPELHKSIVWILENITVLDHTFCVFNHAGRILOHELKPN 539
488 FYKQLGKPIQLSDLESVDPELHKSIVWILENITVLDHTFCVFNHAGRILOHELKPN 547
540 GRNPVTEENKEKVRVLYVNRFRMGTEAQLQKGFNELIPQHLKPFQDKELELIIG 599
548 GKVLQVTEENKEKVRVLYVNRFRMGTEAQLQKGFNELIPQHLKPFQDKELELIIG 607
600 GLDKIDLNDWKSNTRLKHCVADSVNMFVQAVETDEERRALLQFVTGSTRVPLQGFK 659

Db 608 GLDKIDLNDWKSNTRLKHCVADSVNMFVQAVETDEERRALLQFVTGSTRVPLQGFK 667
Qy 660 ALQGSTGAAGPRFLTHLIDANTDNLKPAHCFNRIIDIPPYSEYKLYKLLTAVENTCG 719
Db 668 ALQGSTGAAGPRFLTHLIDANTDNLKPAHCFNRIIDIPPYSEYKLYKLLTAVENTSG 727
Qy 720 FAVE 723
Db 728 FAVE 731

RESULT 3
ID SUFL MOUSE STANDARD; PRT; 619 AA.
AC QSCUN6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
DE (Fragment).
DE (Fragment).
GN SMURF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
STRAN=C57BL/6J; TISSUE=Testis;
MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Matsuda Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Okazaki H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Gustinich M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Brownstein S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maehama J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
EL -!- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BNP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION (BY
CC SIMILARITY).
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AK015264; BAB29770.1; -.
CC HSP; Q13526; IPIN.
CC MGD; MGI:1923038; 4930431E10R1K.
CC GO; GO:0005622; C:intracellular; ISS.
CC GO; GO:0000211; F:protein degradation tagging activity; ISS.
```

GO; GO:0004842; Ubiquitin-protein ligase activity; ISS.
GO; GO:0030154; P:cell differentiation; ISS.
GO; GO:0007398; P:ectoderm development; ISS.
GO; GO:0030514; P:negative regulation of BMP signaling pathway; ISS.
GO; GO:0006464; P:protein modification; ISS.
GO; GO:0006512; P:ubiquitin cycle; ISS.
InterPro; IPR000008; C2
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 2.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 2.
PROSITE; PS00499; C2 DOMAIN 1; PARTIAL.
PROSITE; PS00004; C2 DOMAIN 2; PARTIAL.
PROSITE; PS0237; HECT; 1.
PROSITE; PS01159; WW DOMAIN 1; 1.
PROSITE; PS0020; WW DOMAIN 2; 2.
UB1 conjugation pathway; Ligase; Repeat.
NON TER 1
DOMAIN 125 158 WW 1.
DOMAIN 171 204 WW 2.
DOMAIN 285 619 HECT.
DOMAIN 35 38 POLY-GLY.
SEQUENCE 619 AA; 70902 MW; 4CBE2F6624A7B525 CRC64;
Query Match 83.7%; Score 3249.5; DB 1; Length 619;
Best Local Similarity 97.3%; Pred. No. 5.4e-228;
Matches 605; Conservative 6; Mismatches 3; Gaps 1;
102 DTGYRLDLKLPSTDAVGGIIVSLQTRIGTGGSVDCRGLLENGTVYEDSGP 161
1 DTGYRLDLKLPSTDAVGGIIVSLQTRIGTGGSVDCRGLLENGTVYEDSGP 60
162 RPLSCFEEPAFTDTGAAAGGNCRFVESPSQORLQAQRLNPDVGRSLQTPQNRPH 221
61 RPLSCFEEPAFTDTGAAAGGNCRFVESPSQORLQAQRLNPDVGRSLQTPQNRPH 120
222 GHOSPELPEGYEORTVQGVYFLHTQTGVSTVHDPRIIPRLDLSVNCDELGLPFGWEVR 281
121 GHOSPELPEGYEORTVQGVYFLHTQTGVSTVHDPRIIPRLDLSVNCDELGLPFGWEVR 180
282 STVSGRIYFVDHNNRTTQFTDPLHIMHQCQLKEPSQPLPSEGLSLEDBELPAQRYE 341
181 STVSGRIYFVDHNNRTTQFTDPLHIMHQCQLKEPSQPLPSEGLSLEDBELPAQRYE 240
342 RDLVQKLKVLRLHLSLQOPQAGHCRLEVGREIFEEISYQIMKMRPKDLKGLMVKFRGE 401
241 RDLVQKLKVLRLHLSLQOPQAGHCRLEVGREIFEEISYQIMKMRPKDLKGLMVKFRGE 300
402 EGLDYGGVAREWLYLLCHEMLNPPYGLFQYSTDNIMYLQINPDSSINPDHLSYFHFVGRI 461
301 EGLDYGGVAREWLYLLCHEMLNPPYGLFQYSTDNIMYLQINPDSSINPDHLSYFHFVGRI 360
462 MGLAVFHGYINGGFTVPFYKQLKGRPIQLSDLESVDPELHLSLVILENDITPVLDTFF 521
361 MGLAVFHGYINGGFTVPFYKQLKGRPIQLSDLESVDPELHLSLVILENDITPVLDTFF 420
522 CVEHNAIGRILOHLEKPNVPTVEENKEVYLVVNRFRGTEAQLAKQGFNELI 581
421 CVEHNAIGRILOHLEKPNVPTVEENKEVYLVVNRFRGTEAQLAKQGFNELI 480
582 FOHLKLPFDQKELELIIGGLDKIDLNDWKSNTLRKHCVDASNIIVRFWQAVTFDEERRA 641
481 FOHLKLPFDQKELELIIGGLDKIDLNDWKSNTLRKHCVDASNIIVRFWQAVTFDEERRA 540
642 RLLOFVGTSTVPLQKALOGSTGAGPRLFTHLIDANTONLPAKTCFNRIDIPYE 701
541 RLLOFVGTSTVPLQKALOGSTGAGPRLFTHLIDANTONLPAKTCFNRIDIPYE 597
702 SYEKLYEKLTAVEETCGFAVE 723
598 SYEKLYEKLTAVEETCGFAVE 619

RESULT 4
SUFG2 HUMAN
ID SUFG2_HUMAN STANDARD; PRT; 748 AA.
AC Q9HAU4; Q9H260;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 2 (EC 6.3.2.-) (Ubiquitin--
protein ligase SMURF2) (Smad-specific E3 ubiquitin ligase 2)
DE (hSMURF2).
GN SMURF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF PRO-251--VAL-284 AND
RP GLY-297--LEU-330.
RX PubMed=11163210;
RA Kavsek P., Rasmussen R.K., Causing C.G., Bonni S., Zhu H.,
Thomsen G.H., Wtana J.L.;
RT "Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the
RT TGF-beta receptor for degradation."; Mol. Cell 6:1365-1375 (2000).
RL [2]
RN SEQUENCE FROM N.A., AND MUTAGENESIS OF PRO-251--VAL-284 AND CYS-716.
RP MEDLINE=20538422; PubMed=11016919;
RX Lin X., Liang M., Feng X.-H.;
RT "Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent
RT degradation of Smad2 in transforming growth factor-beta signaling."; J. Biol. Chem. 275:36818-36822 (2000).
RL [3]
RN SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-716.
RP MEDLINE=21107656; PubMed=11158580;
RX Zhang Y., Chang C., Gehling D.J., Hemmati-Brivanlou A., Derynck R.;
RT "Regulation of Smad degradation and activity by Smurf2, an E3
RT ubiquitin ligase."; Proc. Natl. Acad. Sci. U.S.A. 98:974-979 (2001).
CC [1-] FUNCTION: Interacts with SMAD1, SMAD2 and SMAD7 in order to
CC trigger their ubiquitination and proteasome-dependent degradation.
CC Enhances the inhibitory activity of SMAD7 and reduces the
CC transcriptional activity of SMAD2. Coexpression of SMURF2 with
CC SMAD1 results in considerable decrease in steady-state level of
CC SMAD1 protein and a smaller decrease of SMAD2 level.
CC [1-] SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, SMAD6 and SMAD7 but
CC not SMAD4.
CC [1-] SUBCELLULAR LOCATION: Nuclear. Cytoplasmic in the presence of
CC SMAD7.
CC [1-] TISSUE SPECIFICITY: Widely expressed.
CC [1-] DOMAIN: The second and third WW domains are responsible for
CC interaction with R-SMAD (SMAD1, SMAD2 and SMAD3).
CC [1-] SIMILARITY: Contains 1 C2 domain.
CC [1-] SIMILARITY: Contains 3 WW domains.
CC [1-] SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC EMBL; AF310676; AAG45422.1; --
CC EMBL; AF310463; AAG25641.1; --
CC EMBL; AY014180; AAG50421.1; --
CC HSP; Q13526; 1PIN.
CC MIM; 605532; --
CC GO; GO:0004842; F:ubiquitin-protein ligase activity; NAS.
CC GO; GO:0016481; P:negative regulation of transcription; NAS.

GO: GO:0017015; P-regulation of TGFbeta receptor signaling pa. . .; NAS.

InterPro: IPR000008; C2.

InterPro: IPR000569; HECT domain.

InterPro: IPR001202; WW_Rsp5_WWP.

Pfam: PF00168; C2; 1.

Pfam: PF00632; HECT; 1.

Pfam: PF00397; WW; 3.

SMART: SM00239; C2; 1.

SMART: SM00119; HECT; 1.

SMART: SM00456; WW; 3.

PROSITE: PS00499; C2 DOMAIN 1; 1.

PROSITE: PS00004; C2 DOMAIN 2; 1.

PROSITE: PS50237; HECT; 1.

PROSITE: PS01159; WW DOMAIN 1; 1.

PROSITE: PS00020; WW DOMAIN 2; 3.

Ubl conjugation pathway; Ligase; Repeat; Nuclear protein.

DOMAIN 1 98 C2 DOMAIN.

DOMAIN 157 190 WW 1.

DOMAIN 251 284 WW 2.

DOMAIN 297 330 WW 3.

DOMAIN 414 748 HECT.

MUTAGEN 251 284

MISSING: ABOLISHES INTERACTION WITH

SMAD2 AND SMAD7.

MISSING: ABOLISHES INTERACTION WITH

SMAD7.

C->A: LOSS OF ABILITY TO UBIQUITINATE

SMAD2.

C->G: LOSS OF ACTIVITY (LOSS OF ABILITY

TO UBIQUITINATE SMAD1 AND SMAD2 AND NO

DOWN-REGULATION OF SMAD1 AND SMAD2

PROTEIN LEVELS).

G -> R (IN REF. 2).

CONFLICT 6 6

SEQUENCE 748 AA; 86195 MW; 30428443A3755762 CRC64;

Query Match 77.38; Score 3001; DB 1; Length 748;

Best Local Similarity 74.68; Pred No. 7.7e-210;

Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

5 IKRLTVLCARLAKDFFLPDPFAKIVVDGSGQCHSTDTVNTKLDPKWQHLYLVGK 64

12 VKRLTVLCARLAKDFFLPDPFAKIVVDGSGQCHSTDTVNTKLDPKWQHLYLVGK 71

65 TDSITISVWKKHKKOGAGFGLCVRLSLNALSRLKDTGYORLDCIKLPNDTDAVGGQ 124

72 SDSVTISVWKKHKKOGAGFGLCVRLSLNALSRLKDTGYORLDCIKLPNDTDAVGGQ 131

125 IWSLQTRDRIGTGGSVVDCRGLLENE-----GTVT----- 155

132 IVSLQSRDRIGTGGSVVDCRGLLENE-----GTVT----- 191

156 --EDSGRPLSCFMEEPAPYTDSTGAAAGGNCRFVESPSQDRLQORLNPDPVGRSL 213

192 ASEYSSPGRLSCFVDENTPISGNTATG-----QSSDPLAEARRVSRQRHNYM 242

214 QTPQNRPHGQSPPEGYEORTTVQGVFLHTQTGVSTWHPDRIPRLDINSVNCDELQP 273

243 ---SRTHLTPDLPPEGYEORTTVQGVFLHTQTGVSTWHPDRIPRLDINSVNCDELGP 298

274 LPPGWEVSTVSGRIYFVDHNNRTQTDDR-----LHIMNHQCLKERSQPLPISBGS 329

299 LPPGWEIENATGRVYFVDHNNRTQTDDR-----LHIMNHQCLKERSQPLPISBGS 354

330 L---EDELPAQRYERDLVQKLVRLHSLSQPQAGHCRIEVSRREIPEESVQIMQKR 386

355 LCPDDTECLTPRYKRLVQKLVRLHSLSQPQAGHCRIEVSRREIPEESVQIMQKR 414

387 PKDLKRLWKFREGEGLDYGGVAREWLYLICHEMLNPPYGLFQYSTDNIMQLINPDSS 446

415 PKDLKRLWKFREGEGLDYGGVAREWLYLICHEMLNPPYGLFQYSTDNIMQLINPDSS 474

447 INPDHLSYFHFVGRIMGLAVPHGVINGGFTVPYKCLGKPIQLSDLESVPDELHKSIV 506

475 VNPHELSTFHFVGRIMGLAVPHGVINGGFTVPYKCLGKPIQLSDLESVPDELHKSIV 534

QY 507 WLENDITPVLDHTFCVEHNAFGRILOHELKPNGRNVPVTEENKKEVRLYVNRPMRGI 566

DB 535 WLENDITGVLDHTFCVEHNAFGRILOHELKPNGRNVPVTEENKKEVRLYVNRPMRGI 594

QY 567 EAQFLALQKGFNELIPQHLKPFQFQKELELIIGLDKIDLDNWKSNTRLKHCVADSNIVR 626

DB 595 EAQFLALQKGFNEVLPQHLKPFQFQKELELIIGLDKIDLDNWKSNTRLKHCVADSNIVR 654

QY 627 NPWQAVETPDERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLETHILIDANTNLP 686

DB 655 NPWQAVETPDERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLETHILIDANTNLP 711

QY 687 KAHTCFNRIDIPPEVSEYKLYEKLTLTAVEETCGFAVE 723

DB 712 KAHTCFNRIDIPPEVSEYKLYEKLTLTAVEETCGFAVE 748

RESULT 5

PUBI SCHPO STANDARD; PRT; 767 AA.

ID C92452; O14454; AC

DT 01-NOV-1997 (Rel. 35, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Ubiquitin--protein ligase pub1 (EC 6.3.2.-).

GN PUB1 OR SPAC11G7.02.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96205868; PubMed=8635463;

RA Nefsky B., Beach D.;

RT "Publ acts as an E6-AP-like protein ubiquitin ligase in the

RT degradation of cdc25.";

RL EMBO J. 15:1301-1312(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=97340937; PubMed=9197411;

RA Salehi R., Jia Z., Karagiannis J., Young P.G.;

RT "Tolerance of low pH in Schizosaccharomycetes pombe requires a

RT functioning pub1 ubiquitin ligase.";

RL Mol. Gen. Genet. 254:520-528(1997).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy I., Niblett D., Odeil C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Wainman F., Volkart G., Aert R., Robben J., Grymonprez B.,

RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Spakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 "The genome sequence of *Schizosaccharomyces pombe*."
 Nature 415:871-880(2002).
 -!- FUNCTION: REGULATES UBIQUITINATION OF CDC25.
 -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP + diphosphate + protein N-ubiquityllysine.
 -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-thiolester formation.
 -!- SIMILARITY: Contains 1 C2 domain.
 -!- SIMILARITY: Contains 3 WW domains.
 -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.
 -!- SIMILARITY: STRONG, TO YEAST RSP5.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; Y07592; CA368867.1; -;
 EMBL; U66716; AAB07534.1; -;
 EMBL; Z99161; CAB6207.1; -;
 EMBL; U62795; AAB63350.1; -;
 PIR; S66562; S66562.
 PIR; T37545; T37545.
 HSSP; Q13526; 1PIN.
 GeneDB SPombe; SPAC11G7.02; -;
 InterPro; IPR000008; C2.
 InterPro; IPR000569; HECT_domain.
 InterPro; IPR002349; WW.
 InterPro; IPR001202; WW_Rsp5_WWP.
 Pfam; PF00168; C2; 1.
 Pfam; PF00632; HECT; 1.
 Pfam; PF00397; WW; 3.
 PRINTS; PR00403; WWDOMAIN.
 SMART; SM00239; C2; 1.
 SMART; SM00119; HECTC; 1.
 SMART; SM00456; WW; 3.
 PROSITE; PS00499; C2_DOMAIN_1; 1.
 PROSITE; PS50004; C2_DOMAIN_2; 1.
 PROSITE; PS50237; HECT; 1.
 PROSITE; PS01159; WW_DOMAIN_1; 3.
 PROSITE; PS50020; WW_DOMAIN_2; 3.
 US1 conjugation pathway; Ligase; Repeat.
 DOMAIN 17 32
 C2 DOMAIN.
 DOMAIN 211 236
 WW 1.
 DOMAIN 242 247
 POLY-ALA.
 DOMAIN 294 319
 WW 2.
 DOMAIN 351 376
 WW 3.
 DOMAIN 463 767
 HECT.
 BINDING 735
 UBIQUITIN (BY SIMILARITY).
 Q -> K (IN REF. 1).
 CONFLICT 163 609
 MISSING (IN REF. 1).
 CONFLICT 661 661
 T -> K (IN REF. 1).
 CONFLICT 767 AA; 87267 MW; F1455A155EB9ACF7 CRC64;
 SEQUENCE
 Query Match 38.4%; Score 1492.5; DB 1; Length 767;
 Best local similarity 41.1%; Pred. No. 2.1e-100;
 Matches 329; Conservative 118; Mismatches 234; Indels 119; Gaps 19;
 4 SIKRLTVLCAKLNKDFRPPDPFAKTVDPGSGCHSTDTVKNTLDPKNQHYDLYVG 63
 7 SRRIRTVIADGLYKEDVRFPDPFAVTVDGE-QTHTTAIKTLNPNYNEFVNVT 65
 64 KTSITISVNHKKIKHKQAGPLGCVRLLSNAIRLKDVTGVRQL--DLCKLNPSDITAV 121
 66 DNSTIATQVDFDKF-KKKGQGLVNLRVGVDLDAIGGDEMLTRDLKKS--ENTTV 122
 122 RGQIVVSLQ-----TDRIGTGSVVDRCGLLENESGTVYED 157

Db 123 HGKIIINLITTAQSTLQVPSAASGARTQRTSITNDPQSSQSSSVSRNPASSRAGSPTE 182
 QY 158 SGPG-----RPLSCFMBE-----PAPY-----TDSTG-----AAA 182
 Db 183 NAPAASPASSEPRTFSEFQYGRLPFGWERRTDNLGRTVYVDHNRSTTWIRPNLSSVA 242
 QY 183 GGGCRPFVESPS-----QDQRLQARLNRNDVGRSLQTPQNRPHGHSPELPGYEQ 234
 Db 243 GAAAEHLSSASSANVTGQVPSSSNNAARTEASVLTSTNATTAG-----SGELPGWQ 296
 QY 235 RTTVQGVYFLHTQTGVSTVHWDPRIPRDLNSVN-----CDELGPLPGWVRST 283
 Db 297 RYTPGGRPYFVDHNRTRTTTWDPRRQYIRSYGPNNAITQQQPVSQLGPLSGWEMRLT 356
 QY 284 VSGRTYFVDHNRTRTTQTDRLHHVNHQCLKEPSQPLPSEGSLEDEELPAQRYERD 343
 Db 357 NIARVYFVDHNRKTTTWDPR-----LPS--SL-DQVNP--QYKRD 392
 QY 344 LVQKLKVLRLHLSLQQAQGHCRIEYSREIFEESYRQIMQMRPKDLKRLMYKFRGEG 403
 Db 393 FRKLIYFLSQPAL-HPLPGQCHIKVRNHFEDSYAEIMRQSATDLKRLMIKPDGEG 451
 QY 404 LDYGGVAREWLYLICHEMLNPPYCLFOYSTDNLYMLQINPDSSINPDHLSYEFVGRING 463
 Db 452 LDYGLSREYFFLLSHENFPFYCLFEYSVDNTYQINPHSGINFEHLNYFAFGRVIG 511
 QY 464 LAVFHGHYINGFTVPFYKOLLGKPIQLSDLESVDPELHKSLVWILENDITPVLDHTFCV 523
 Db 512 LAIFRRFVDAFFVVFYKMLQKVLQDMESMDAEYRSVLWILDNDITGVLDLTFVS 571
 QY 524 EHNAFGRILQHELPNGRNPVTENKKEVRLVYVWRPFRGTEAQFLALQKFNELIQ 583
 Db 572 EDNCFGEVVTIDLPNGRNIETVEENKREYVDLVTVMRIQRIEEOFNAPHEGFSLEIPQ 631
 QY 584 HLLKPFDOKELELIIGGLDKIDLNDKSNTRLKHCVADSNIVRFWQAVETFEERRARL 643
 Db 632 ELINVFDERELELLIGGISEIDMEDWKKHTDYSSENDOIKWFELMDENSEKKSRL 691
 QY 644 LQFTVGTSTRVPLQGFKALQGSTGAAGPRLFTIHLIDANTNLPAKJTCNRIIDIPYESY 703
 Db 692 LQFTTGTSTRVNGFKDLQSGD---GPRKFTIEKA-GEPNKLPKAHTCFNRLDLPYTSK 747
 QY 704 EKLVEKLLTAVEETCGFAVE 723
 Db 748 KDLHKLSLIAVEETIGFGQE 767
 RESULT 6
 RSP5 YEAST
 ID RSP5 YEAST STANDARD; PRT; 809 AA.
 AC P39940.
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-protein ligase RSP5 (EC 6.3.2.-)
 GN RSP5 OR NP11 OR MDP1 OR YER125W OR SYGP-ORF41.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; *Saccharomycotina*; *Saccharomycetes*;
 OC *Saccharomycetales*; *Saccharomycetaceae*; *Saccharomyces*.
 CX NCBI_TaxID=4932;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX PubMed=9169868;
 RA Dietrich P.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
 Huntcke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V.";

Nature 387:78-81 (1997).
[2] IDENTIFICATION.
Winston F.;
Unpublished observations (FEB-1993).
[3]
CHARACTERIZATION.
STRAIN=Sigma 1278B;
MEDLINE=96154942; PubMed=8596462;
Hein C., Springsael J.-Y., Volland C., Hagenauer-Teaplis R., Andre B.;
"Nfil, an essential yeast gene involved in induced degradation of Gap1
and Fur4 permeases, encodes the Rsp5 ubiquitin-protein ligase.";
Mol. Microbiol. 18:77-87 (1995).
[4]
FUNCTION.
MEDLINE=95223981; PubMed=7705685;
Huibregtse J.M., Scheffner M., Beaudenon S., Howley P.M.;
"A family of proteins structurally and functionally related to the
E6-AP ubiquitin-protein ligase.";
Proc. Natl. Acad. Sci. U.S.A. 92:2563-2567 (1995).
[5]
ERRATUM.
MEDLINE=95281634; PubMed=7761480;
Huibregtse J.M., Scheffner M., Beaudenon S., Howley P.M.;
Proc. Natl. Acad. Sci. U.S.A. 92:5249-5249 (1995).
[6]
CHARACTERIZATION.
STRAIN=FV56;
MEDLINE=99077972; PubMed=9858558;
Wang G., Yang J., Huibregtse J.M.;
"Functional domains of the rsp5 ubiquitin-protein ligase.";
Mol. Cell. Biol. 13:342-352 (1993).
[7]
-!- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM
AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND
THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES.
REQUIRED FOR UBIQUITINATION AND THEREFORE DEGRADATION OF SEVERAL
CELL SURFACE PROTEINS LIKE GAP1, FUR4, MAL61 AND STE2. ALSO ACTS
ON RBP1.
-!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (Potential).
-!- PTM: THE UBIQUITINATION APPEARS TO BE THE RESULT OF AN
INTRAMOLECULAR TRANSFER OF UBIQUITIN.
-!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
thioester formation.
-!- SIMILARITY: Contains 1 C2 domain.
-!- SIMILARITY: Contains 3 WW domains.
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-!- SIMILARITY: STRONG, TO S.POMBE PUBL.
-!- SIMILARITY: TO YEAST YKL010C.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: U18916; AAC03223.1; -;
PIR: S43217; S43217.
HSSP: Q13526; IPIF.
SGP: S0000927; RSPS.
GO: GO:0000151; C:ubiquitin ligase complex; IDA.
GO: GO:0004842; F:ubiquitin-protein ligase activity; IDA.
GO: GO:0006333; P:chromatin assembly/disassembly; IMP.
GO: GO:0006513; P:protein monoubiquitination; IDA.
GO: GO:0000209; P:protein polyubiquitination; IDA.
InterPro: IPR000008; C2.
InterPro: IPR000569; HECT_domain.
InterPro: IPR002349; WW.
Pfam: PF00168; C2.1.
Pfam: PF00632; HECT; 1.

DR Pfam: PF00397; WW; 3
DR PRINTS; PRO0403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS00237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS00020; WW_DOMAIN_2; 3.
KW Ubl conjugation pathway; Ligase; Repeat.
FT DOMAIN 1 88
FT DOMAIN 229 262
FT DOMAIN 315 322
FT DOMAIN 331 364
FT DOMAIN 387 420
FT DOMAIN 705 809
FT BINDING 777 777
FT VARIANT 733 733
FT MUTAGEN 777 777
FT SEQUENCE 809 AA; 91816 MW; 6F1836384479E70F CRC64;
Query Match 35.5%; Score 1377; DB 1; Length 809;
Best Local Similarity 37.0%; Pred. No. 5.5e-92;
Matches 313; Conservative 131; Mismatches 238; Indels 164; Gaps 24;
QY 3 SSIKRLTVLCAKMLAKKDFRPPFAKIVVDGSGOCHSTDTVKNTLDPKWNCHYDL-Y 61
DB 3 SSIIVKLV--VAESLYKRDVFRSPDPFVAVIDGY-QTKSTSAKKTLYNFWNETFFDD 59
QY 62 VKGTDSITISVNNHKKIKKOGAGFLGCVRL-LSNAISRLKD-----TGQVR-----LDL 110
DB 60 INENSILITIQVEDQKFF-KKKDQGFVGVVAVGVGLHLEDATSSGRPREETITRDL 118
QY 111 CKLNPSDITDAVSGQIVV-----SLQTRDRIGTSGSVVDCRGLL 148
DB 119 KKN--DGMVSGRLIVLVLSKLPSSPHSQAPSGHTASSNTSTTRTNGHSTSS---T 173
QY 149 ENEGTVYEDSGFRLSCFMEEPAP-----YDSTGAAA-----182
DB 174 RNHSTSHPSRGTAQAVESTLQSGTTAATNTATTHRSSTNSTSSATQVSSPEDQYGRLLP 233
QY 183 -----GGGNCRFVE-----SPSQDQLQAORLNDPVRGSLQTPQR-----PHG 222
DB 234 GMERPTDNFRYYVDHNTRTTWKRPFLDQ-TEABRGNQLNANTELERRORHRTLPGG 292
QY 223 HQS-----PELPEGYEORTTVQGVYFLHT 247
DB 293 SSDNSVTYQVGGGNIIPVNGAAAAAFAATGTTSGELPSCGEGRFTPEGRAYFVDH 352
QY 248 QGVSTWHDPRIPDLNSVN-----CDEGLPLPGWEVSTVSGRIYFVDHNRRT 297
DB 353 NTRTTTVDVPRQQYIRTYGPTNTTIOQQPVSQLPSPGWEMELTNTARYFYVDHNTKT 412
QY 298 TQPTDRLHMHNHQOLKERSQPLPSEGSLEDELPQORYERDLVQKLKLRHLSL 357
DB 413 TWDDPR-----LPS--SL-DQVNP--QYKRDPRKVIYFRSQPAL 448
QY 358 QPQAGHCRIEVSREIEFESYRQIMKVRPKDLKRLMKVPRGEGLDYGVAREWLYLL 417
DB 449 -RLPGQCHIKVRRKNIFEDAYQEIIMRQTPEDLKKRLMKPDGEGLDYGVAREWLYLL 507
QY 418 CHEMLNPPYGLFOYSTDNILYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHGHYNGFT 477
DB 508 SHEMFNFFCLFEYSAYNYTIQINPNSGINPEHLNFKFICRVVGLGVFHRRLDFAFFV 567
QY 478 VPFYKQLGKPIQLSDLESVDPELHSLVWILENDITPVLDPHTFCVHNFAFRILQHELK 537
DB 568 GALYKMWLKKKVVLDQMEGVDAEVVNSLNWMLNSIDGLVLDLTFSADDERFGEVTVDLK 627
QY 538 PNGRNVPTVEENKKEYVRLYVNRFRMRCIEAQFLAQKGFNELIPQHLLKPFQCKELELI 597

[illegible]

557 KPFOKELELIIGLKDIDNDKSNTRLKH-CVADSNIWVFWQAVETEDERRARLLO 645
 793 KIFDENELELMCGLDVNDWREHTKYXNGYSANFQVQWFKAVLMDWSEKIRLLO 852
 646 FVTGSTRVPLGFKALQOSTGAAPRFTTHLIDANTDNLPKAHTCFNRIIDIPPVESYEK 705
 853 FVTGSTRVPMNGFAELYGSN--GQSFTEQW-GTPEKLPRAHTCFNRLDIPPVESFEE 908
 706 LYEKLLTAVEETCGF 720
 903 LWDLQMAIENTQGF 923

ULT 8

4_MOUSE

SEQUENCE STANDARD; PRT; 957 AA.

01-NOV-1995 (Rel. 32, Created)
 01-OCT-1996 (Rel. 34, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 NEDD-4 protein (EC 5.3.2.-) (Fragment).
 NEDD4 OR NEDD-4 OR NEDD4A.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Embryo.
 MEDLINE=92328780; PubMed=1378265;
 Kumar S.; Tomooka Y.; Noda M.;
 "Identification of a set of genes with developmentally down-regulated
 expression in the mouse brain."
 Biochem. Biophys. Res. Commun. 185:1155-1161(1992).
 [2]

REVISIONS.

Kumar S.;
 Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM
 AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND
 THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES (BY
 SIMILARITY).
 -!- FUNCTION: INVOLVED IN THE EMBRYONIC DEVELOPMENT AND
 DIFFERENTIATION OF THE CENTRAL NERVOUS SYSTEM.
 -!- TISSUE SPECIFICITY: BRAIN.
 -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 thiolester formation.
 -!- SIMILARITY: Contains 1 C2 domain.
 -!- SIMILARITY: Contains 3 WW domains.
 -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; D85414; BAA12803.1; -
 HSSP; Q13526; 1PIN.
 MSD; MGI:97297; Nedd4.
 GO; GO:0005829; Cytozol; IDA.
 GO; GO:0000151; C:ubiquitin ligase complex; IPI.
 GO; GO:0005515; F:protein binding activity; IPI.
 GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
 InterPro; IPR000008; C2.
 InterPro; IPR000569; HECT domain.
 InterPro; IPR001202; WW_Rsp5_WWP.
 Pfam; PF00168; C2, 1.
 Pfam; PF00632; HECT; 1.

DR Pfam; PF00397; WW; 3.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00119; HECT; 1.
 DR SMART; SMO0456; WW; 3.
 DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 DR PROSITE; PS00004; C2_DOMAIN_2; 1.
 DR PROSITE; PS0237; HECT; 1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 3.
 DR PROSITE; PS00020; WW_DOMAIN_2; 3.
 KW Ubl conjugation pathway; Ligase; Repeat.
 FT NON TER 1
 FT DOMAIN 150 236 C2 DOMAIN.
 FT DOMAIN 319 352 WW 1.
 FT DOMAIN 475 508 WW 2.
 FT DOMAIN 530 563 WW 3.
 FT DOMAIN 851 957 HECT.
 FT BINDING 924 924 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 957 AA; 109967 MW; 79D5F1DF0187F28 CRC64;
 Query Match 31.5%; Score 1225; DB 1; Length 957;
 Best Local Similarity 33.6%; Pred. No. 7.5e-81;
 Matches 291; Conservative 123; Mismatches 238; Indels 214; Gaps 23;
 QY 7 IRLTVLCAKYLAKDFRLPDPFAKIV----VDGSGQCHSTDTVKNTLDPKGNQHY---- 58
 DB 150 VRKVIAGIAGLAKDILGASDPVRYVTLVDYPMGILTSVQTKIKSLNPKWNEILFRV 209
 QY 59 ---DLVYVKTD--SITISVWN-----HKKIKKQG 93
 DB 210 LPQHRHLPEVFDENRLTRDDFLGQVDVPLYPLETPNPRMERPYTFKDFVLHPRSHKSRV 269
 QY 84 AGFLGCVRLLSNAISRLKDT----- 103
 DB 270 KGYL-----RLKMTYLPKNGSEDENADQAELEPGVWVLDQDPAATHLPHPEP 318
 QY 104 ---GY-QRLDLC-----KLNPSDTDAVRGQIVVSLQTRDRI 135
 DB 319 SPLPFGHEERQDVLGRYYVNHESRRTQMKRSPDDDLTDEDNDMLQAOAQAFTTRQI 378
 QY 136 GTGGSVVDCRGLLEN-----EGTVYEDSGGCRPLSCFME-----EP 171
 DB 379 SEDVDGPDNRESPENWEIVREDENTEYSQAVSQSPSGHIDVQTHLAEFNTRLAVCGNP 438
 QY 172 A---PYTDTGAAAGG--NCRVESPSQDRLQAOQLRANPDVRGSLQTPQNRPHGHQSP 226
 DB 439 ATSPVTSSNHSRGGSLQTCIFEQPTLPVLLPT-----SS 475
 QY 227 ELPEGYQRTTVGQVYFLHTQTGVSTWHD-----RIPRDL-NSVNCDELGLPLPG 277
 DB 476 GLPPGWEKQDDGRSYVVDHNSKTTTWSKPTQDDPSRKIPAHLRGKTDSDNDGLGPLPFG 535
 QY 278 WEVRSTVSGRIYFVDHNNRTTQTPDLRHNNHQCQLAKEPSQPLPSEGSLEDEELPA 337
 DB 536 WEERTHTDGRVFFINHNKTKTQWEDPRLQNV-----AITGPAVP----- 574
 QY 338 QRYRDLVQKLKVLRLHLSLQFPQAGHCRIEVSREIFEEYSRQIMMMPKD-LKKRLMY 396
 DB 575 --YSDYKRYKRYEFERRKLLKQTDIPNKFEMKLARANILEDYSYRIRMGVXRADILKARLWI 632
 QY 397 KFRREGLDYGVAREWMLYLICHEMLNPYGLFQYSTDNIMQLINPDSSI-NPDHLSYF 455
 DB 633 EFDGKGLDYGVAREWFFLIKEMFNFPYGLFYSATDNTYLIQINPNSGLCNEDHLSYF 692
 QY 456 HFVGRIMGLAVFHGHYINGGTFVPFYKQLGKGIQSLDSLESDVPELHKSILVWLINDITP 515
 DB 693 KFIQGVAGMAYVHGKLLDGGFIRFPYKWLQKLTLDHMSVDSEYSSILRWLENDPTE 752
 QY 516 VLDHTFCVHNNAFGRILOHEIKNGRNVPTTEKNKYVRLVYNNFMFGIEAQFLALQK 575
 DB 753 -LDLRFIIDEELFGQTHQELKTGSGSEIVVTNKNKBYLYLVQWRFVARIQKMAAFKE 811
 QY 576 GFNELIPQHLAKPFDQKELELIIGGLDKIDLNDKWNTRLKHCA-DSNIVRWFQAVET 934

"Mutation analysis of UBE3A in Angelman syndrome patients."; Am. J. Hum. Genet. 62:1353-1360(1998).

CHARACTERIZATION

MEDLINE=98351563; PubMed=9688277;

Nuber U., Schwarz S.B., Scheffner M.;

"The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as its own substrate.";

Eur. J. Biochem. 254:643-649(1998).

!- FUNCTION: INTERACTS WITH THE E6 PROTEIN OF THE CANCER-ASSOCIATED HUMAN PAPILLOMAVIRUS TYPES 16 AND 18. THE E6/E6-AP COMPLEX BINDS TO AND TARGETS THE P53 TUMOR SUPPRESSOR PROTEIN FOR UBIQUITIN-MEDIATED PROTEOLYSIS. IT IS AN E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES. IT CAN TARGET ITSELF FOR UBIQUITINATION IN VITRO AND EFFICIENTLY PROMOTES ITS OWN DEGRADATION IN VIVO. IT APPEARS THAT ONLY UNMODIFIED E6-AP MOLECULES CAN BIND EFFICIENTLY TO P53 IN THE PRESENCE OF THE HPV E6 ONCOPROTEIN.

!- SUBCELLULAR LOCATION: Nuclear (Probable).

!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=3;

Name=I; IsoId=Q05086-1; Sequence=VSP_006705;

Name=II; IsoId=Q05086-2; Sequence=VSP_006705;

Name=III; IsoId=Q05086-3; Sequence=VSP_006706;

!- DISEASE: DEFECTS IN UBE3A ARE A CAUSE OF ANGELMAN SYNDROME (AS), ALSO KNOWN AS 'HAPPY PUPPET SYNDROME'. IT IS CHARACTERIZED BY FEATURES OF SEVERE MOTOR AND INTELLECTUAL RETARDATION,

MICROCEPHALY, ATAXIA, FREQUENT JERKY LIMB MOVEMENTS AND FLAPPING OF THE ARMS AND HANDS, HYPOTONIA, HYPERACTIVITY, HYPOPIGMENTATION, SEIZURES, ABSENCE OF SPEECH, FREQUENT SMILING AND EPISODES OF PAROXYSMAL LAUGHTER, AND AN UNUSUAL FACIES CHARACTERIZED BY MACROSTOMIA, A LARGE MANDIBLE AND OPEN-MOUTHED EXPRESSION, A GREAT PROPENSITY FOR PROTRUDING THE TONGUE ('TONGUE THRUSTING'), AND AN OCCIPITAL GROOVE.

!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOLESTER FORMATION.

!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.

!- SIMILARITY: SOME, IN THE C-TERMINAL HALF, TO RAT 100 kDa PROTEIN. THIS SWISS-PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X98021; CAA66653.1; JOINED.
EMBL; X98027; CAA66653.1; JOINED.
EMBL; X98022; CAA66653.1; JOINED.
EMBL; X98023; CAA66653.1; JOINED.
EMBL; X98024; CAA66653.1; JOINED.
EMBL; X98025; CAA66653.1; JOINED.
EMBL; X98026; CAA66653.1; JOINED.
EMBL; X98028; CAA66653.1; JOINED.
EMBL; X98029; CAA66653.1; JOINED.
EMBL; X98030; CAA66653.1; JOINED.
EMBL; X98033; CAA66653.1; JOINED.
EMBL; X98031; CAA66653.1; JOINED.
EMBL; X98032; CAA66653.1; JOINED.
EMBL; X98033; CAA66653.1; JOINED.
EMBL; X98034; CAA66653.1; JOINED.
EMBL; X98035; CAA66653.1; JOINED.
EMBL; X98036; CAA66653.1; JOINED.
EMBL; X98037; CAA66653.1; JOINED.
EMBL; X98038; CAA66653.1; JOINED.
EMBL; X98039; CAA66653.1; JOINED.
EMBL; X98040; CAA66653.1; JOINED.
EMBL; X98041; CAA66653.1; JOINED.
EMBL; X98042; CAA66653.1; JOINED.
EMBL; X98043; CAA66653.1; JOINED.
EMBL; X98044; CAA66653.1; JOINED.
EMBL; X98045; CAA66653.1; JOINED.
EMBL; X98046; CAA66653.1; JOINED.
EMBL; X98047; CAA66653.1; JOINED.
EMBL; X98048; CAA66653.1; JOINED.
EMBL; X98049; CAA66653.1; JOINED.
EMBL; X98050; CAA66653.1; JOINED.
EMBL; X98051; CAA66653.1; JOINED.
EMBL; X98052; CAA66653.1; JOINED.
EMBL; X98053; CAA66653.1; JOINED.
EMBL; X98054; CAA66653.1; JOINED.
EMBL; X98055; CAA66653.1; JOINED.
EMBL; X98056; CAA66653.1; JOINED.
EMBL; X98057; CAA66653.1; JOINED.
EMBL; X98058; CAA66653.1; JOINED.
EMBL; X98059; CAA66653.1; JOINED.
EMBL; X98060; CAA66653.1; JOINED.
EMBL; X98061; CAA66653.1; JOINED.
EMBL; X98062; CAA66653.1; JOINED.
EMBL; X98063; CAA66653.1; JOINED.
EMBL; X98064; CAA66653.1; JOINED.
EMBL; X98065; CAA66653.1; JOINED.
EMBL; X98066; CAA66653.1; JOINED.
EMBL; X98067; CAA66653.1; JOINED.
EMBL; X98068; CAA66653.1; JOINED.
EMBL; X98069; CAA66653.1; JOINED.
EMBL; X98070; CAA66653.1; JOINED.
EMBL; X98071; CAA66653.1; JOINED.
EMBL; X98072; CAA66653.1; JOINED.
EMBL; X98073; CAA66653.1; JOINED.
EMBL; X98074; CAA66653.1; JOINED.
EMBL; X98075; CAA66653.1; JOINED.
EMBL; X98076; CAA66653.1; JOINED.
EMBL; X98077; CAA66653.1; JOINED.
EMBL; X98078; CAA66653.1; JOINED.
EMBL; X98079; CAA66653.1; JOINED.
EMBL; X98080; CAA66653.1; JOINED.
EMBL; X98081; CAA66653.1; JOINED.
EMBL; X98082; CAA66653.1; JOINED.
EMBL; X98083; CAA66653.1; JOINED.
EMBL; X98084; CAA66653.1; JOINED.
EMBL; X98085; CAA66653.1; JOINED.
EMBL; X98086; CAA66653.1; JOINED.
EMBL; X98087; CAA66653.1; JOINED.
EMBL; X98088; CAA66653.1; JOINED.
EMBL; X98089; CAA66653.1; JOINED.
EMBL; X98090; CAA66653.1; JOINED.
EMBL; X98091; CAA66653.1; JOINED.
EMBL; X98092; CAA66653.1; JOINED.
EMBL; X98093; CAA66653.1; JOINED.
EMBL; X98094; CAA66653.1; JOINED.
EMBL; X98095; CAA66653.1; JOINED.
EMBL; X98096; CAA66653.1; JOINED.
EMBL; X98097; CAA66653.1; JOINED.
EMBL; X98098; CAA66653.1; JOINED.
EMBL; X98099; CAA66653.1; JOINED.
EMBL; X98100; CAA66653.1; JOINED.
EMBL; X98101; CAA66653.1; JOINED.
EMBL; X98102; CAA66653.1; JOINED.
EMBL; X98103; CAA66653.1; JOINED.
EMBL; X98104; CAA66653.1; JOINED.
EMBL; X98105; CAA66653.1; JOINED.
EMBL; X98106; CAA66653.1; JOINED.
EMBL; X98107; CAA66653.1; JOINED.
EMBL; X98108; CAA66653.1; JOINED.
EMBL; X98109; CAA66653.1; JOINED.
EMBL; X98110; CAA66653.1; JOINED.
EMBL; X98111; CAA66653.1; JOINED.
EMBL; X98112; CAA66653.1; JOINED.
EMBL; X98113; CAA66653.1; JOINED.
EMBL; X98114; CAA66653.1; JOINED.
EMBL; X98115; CAA66653.1; JOINED.
EMBL; X98116; CAA66653.1; JOINED.
EMBL; X98117; CAA66653.1; JOINED.
EMBL; X98118; CAA66653.1; JOINED.
EMBL; X98119; CAA66653.1; JOINED.
EMBL; X98120; CAA66653.1; JOINED.
EMBL; X98121; CAA66653.1; JOINED.
EMBL; X98122; CAA66653.1; JOINED.
EMBL; X98123; CAA66653.1; JOINED.
EMBL; X98124; CAA66653.1; JOINED.
EMBL; X98125; CAA66653.1; JOINED.
EMBL; X98126; CAA66653.1; JOINED.
EMBL; X98127; CAA66653.1; JOINED.
EMBL; X98128; CAA66653.1; JOINED.
EMBL; X98129; CAA66653.1; JOINED.
EMBL; X98130; CAA66653.1; JOINED.
EMBL; X98131; CAA66653.1; JOINED.
EMBL; X98132; CAA66653.1; JOINED.
EMBL; X98133; CAA66653.1; JOINED.
EMBL; X98134; CAA66653.1; JOINED.
EMBL; X98135; CAA66653.1; JOINED.
EMBL; X98136; CAA66653.1; JOINED.
EMBL; X98137; CAA66653.1; JOINED.
EMBL; X98138; CAA66653.1; JOINED.
EMBL; X98139; CAA66653.1; JOINED.
EMBL; X98140; CAA66653.1; JOINED.
EMBL; X98141; CAA66653.1; JOINED.
EMBL; X98142; CAA66653.1; JOINED.
EMBL; X98143; CAA66653.1; JOINED.
EMBL; X98144; CAA66653.1; JOINED.
EMBL; X98145; CAA66653.1; JOINED.
EMBL; X98146; CAA66653.1; JOINED.
EMBL; X98147; CAA66653.1; JOINED.
EMBL; X98148; CAA66653.1; JOINED.
EMBL; X98149; CAA66653.1; JOINED.
EMBL; X98150; CAA66653.1; JOINED.
EMBL; X98151; CAA66653.1; JOINED.
EMBL; X98152; CAA66653.1; JOINED.
EMBL; X98153; CAA66653.1; JOINED.
EMBL; X98154; CAA66653.1; JOINED.
EMBL; X98155; CAA66653.1; JOINED.
EMBL; X98156; CAA66653.1; JOINED.
EMBL; X98157; CAA66653.1; JOINED.
EMBL; X98158; CAA66653.1; JOINED.
EMBL; X98159; CAA66653.1; JOINED.
EMBL; X98160; CAA66653.1; JOINED.
EMBL; X98161; CAA66653.1; JOINED.
EMBL; X98162; CAA66653.1; JOINED.
EMBL; X98163; CAA66653.1; JOINED.
EMBL; X98164; CAA66653.1; JOINED.
EMBL; X98165; CAA66653.1; JOINED.
EMBL; X98166; CAA66653.1; JOINED.
EMBL; X98167; CAA66653.1; JOINED.
EMBL; X98168; CAA66653.1; JOINED.
EMBL; X98169; CAA66653.1; JOINED.
EMBL; X98170; CAA66653.1; JOINED.
EMBL; X98171; CAA66653.1; JOINED.
EMBL; X98172; CAA66653.1; JOINED.
EMBL; X98173; CAA66653.1; JOINED.
EMBL; X98174; CAA66653.1; JOINED.
EMBL; X98175; CAA66653.1; JOINED.
EMBL; X98176; CAA66653.1; JOINED.
EMBL; X98177; CAA66653.1; JOINED.
EMBL; X98178; CAA66653.1; JOINED.
EMBL; X98179; CAA66653.1; JOINED.
EMBL; X98180; CAA66653.1; JOINED.
EMBL; X98181; CAA66653.1; JOINED.
EMBL; X98182; CAA66653.1; JOINED.
EMBL; X98183; CAA66653.1; JOINED.
EMBL; X98184; CAA66653.1; JOINED.
EMBL; X98185; CAA66653.1; JOINED.
EMBL; X98186; CAA66653.1; JOINED.
EMBL; X98187; CAA66653.1; JOINED.
EMBL; X98188; CAA66653.1; JOINED.
EMBL; X98189; CAA66653.1; JOINED.
EMBL; X98190; CAA66653.1; JOINED.
EMBL; X98191; CAA66653.1; JOINED.
EMBL; X98192; CAA66653.1; JOINED.
EMBL; X98193; CAA66653.1; JOINED.
EMBL; X98194; CAA66653.1; JOINED.
EMBL; X98195; CAA66653.1; JOINED.
EMBL; X98196; CAA66653.1; JOINED.
EMBL; X98197; CAA66653.1; JOINED.
EMBL; X98198; CAA66653.1; JOINED.
EMBL; X98199; CAA66653.1; JOINED.
EMBL; X98200; CAA66653.1; JOINED.
EMBL; X98201; CAA66653.1; JOINED.
EMBL; X98202; CAA66653.1; JOINED.
EMBL; X98203; CAA66653.1; JOINED.
EMBL; X98204; CAA66653.1; JOINED.
EMBL; X98205; CAA66653.1; JOINED.
EMBL; X98206; CAA66653.1; JOINED.
EMBL; X98207; CAA66653.1; JOINED.
EMBL; X98208; CAA66653.1; JOINED.
EMBL; X98209; CAA66653.1; JOINED.
EMBL; X98210; CAA66653.1; JOINED.
EMBL; X98211; CAA66653.1; JOINED.
EMBL; X98212; CAA66653.1; JOINED.
EMBL; X98213; CAA66653.1; JOINED.
EMBL; X98214; CAA66653.1; JOINED.
EMBL; X98215; CAA66653.1; JOINED.
EMBL; X98216; CAA66653.1; JOINED.
EMBL; X98217; CAA66653.1; JOINED.
EMBL; X98218; CAA66653.1; JOINED.
EMBL; X98219; CAA66653.1; JOINED.
EMBL; X98220; CAA66653.1; JOINED.
EMBL; X98221; CAA66653.1; JOINED.
EMBL; X98222; CAA66653.1; JOINED.
EMBL; X98223; CAA66653.1; JOINED.
EMBL; X98224; CAA66653.1; JOINED.
EMBL; X98225; CAA66653.1; JOINED.
EMBL; X98226; CAA66653.1; JOINED.
EMBL; X98227; CAA66653.1; JOINED.
EMBL; X98228; CAA66653.1; JOINED.
EMBL; X98229; CAA66653.1; JOINED.
EMBL; X98230; CAA66653.1; JOINED.
EMBL; X98231; CAA66653.1; JOINED.
EMBL; X98232; CAA66653.1; JOINED.
EMBL; X98233; CAA66653.1; JOINED.
EMBL; X98234; CAA66653.1; JOINED.
EMBL; X98235; CAA66653.1; JOINED.
EMBL; X98236; CAA66653.1; JOINED.
EMBL; X98237; CAA66653.1; JOINED.
EMBL; X98238; CAA66653.1; JOINED.
EMBL; X98239; CAA66653.1; JOINED.
EMBL; X98240; CAA66653.1; JOINED.
EMBL; X98241; CAA66653.1; JOINED.
EMBL; X98242; CAA66653.1; JOINED.
EMBL; X98243; CAA66653.1; JOINED.
EMBL; X98244; CAA66653.1; JOINED.
EMBL; X98245; CAA66653.1; JOINED.
EMBL; X98246; CAA66653.1; JOINED.
EMBL; X98247; CAA66653.1; JOINED.
EMBL; X98248; CAA66653.1; JOINED.
EMBL; X98249; CAA66653.1; JOINED.
EMBL; X98250; CAA66653.1; JOINED.
EMBL; X98251; CAA66653.1; JOINED.
EMBL; X98252; CAA66653.1; JOINED.
EMBL; X98253; CAA66653.1; JOINED.
EMBL; X98254; CAA66653.1; JOINED.
EMBL; X98255; CAA66653.1; JOINED.
EMBL; X98256; CAA66653.1; JOINED.
EMBL; X98257; CAA66653.1; JOINED.
EMBL; X98258; CAA66653.1; JOINED.
EMBL; X98259; CAA66653.1; JOINED.
EMBL; X98260; CAA66653.1; JOINED.
EMBL; X98261; CAA66653.1; JOINED.
EMBL; X98262; CAA66653.1; JOINED.
EMBL; X98263; CAA66653.1; JOINED.
EMBL; X98264; CAA66653.1; JOINED.
EMBL; X98265; CAA66653.1; JOINED.
EMBL; X98266; CAA66653.1; JOINED.
EMBL; X98267; CAA66653.1; JOINED.
EMBL; X98268; CAA66653.1; JOINED.
EMBL; X98269; CAA66653.1; JOINED.
EMBL; X98270; CAA66653.1; JOINED.
EMBL; X98271; CAA66653.1; JOINED.
EMBL; X98272; CAA66653.1; JOINED.
EMBL; X98273; CAA66653.1; JOINED.
EMBL; X98274; CAA66653.1; JOINED.
EMBL; X98275; CAA66653.1; JOINED.
EMBL; X98276; CAA66653.1; JOINED.
EMBL; X98277; CAA66653.1; JOINED.
EMBL; X98278; CAA66653.1; JOINED.
EMBL; X98279; CAA66653.1; JOINED.
EMBL; X98280; CAA66653.1; JOINED.
EMBL; X98281; CAA66653.1; JOINED.
EMBL; X98282; CAA66653.1; JOINED.
EMBL; X98283; CAA66653.1; JOINED.
EMBL; X98284; CAA66653.1; JOINED.
EMBL; X98285; CAA66653.1; JOINED.
EMBL; X98286; CAA66653.1; JOINED.
EMBL; X98287; CAA66653.1; JOINED.
EMBL; X98288; CAA66653.1; JOINED.
EMBL; X98289; CAA66653.1; JOINED.
EMBL; X98290; CAA66653.1; JOINED.
EMBL; X98291; CAA66653.1; JOINED.
EMBL; X98292; CAA66653.1; JOINED.
EMBL; X98293; CAA66653.1; JOINED.
EMBL; X98294; CAA66653.1; JOINED.
EMBL; X98295; CAA66653.1; JOINED.
EMBL; X98296; CAA66653.1; JOINED.
EMBL; X98297; CAA66653.1; JOINED.
EMBL; X98298; CAA66653.1; JOINED.
EMBL; X98299; CAA66653.1; JOINED.
EMBL; X98300; CAA66653.1; JOINED.
EMBL; X98301; CAA66653.1; JOINED.
EMBL; X98302; CAA66653.1; JOINED.
EMBL; X98303; CAA66653.1; JOINED.
EMBL; X98304; CAA66653.1; JOINED.
EMBL; X98305; CAA66653.1; JOINED.
EMBL; X98306; CAA66653.1; JOINED.
EMBL; X98307; CAA66653.1; JOINED.
EMBL; X98308; CAA66653.1; JOINED.
EMBL; X98309; CAA66653.1; JOINED.
EMBL; X98310; CAA66653.1; JOINED.
EMBL; X98311; CAA66653.1; JOINED.
EMBL; X98312; CAA66653.1; JOINED.
EMBL; X98313; CAA66653.1; JOINED.
EMBL; X98314; CAA66653.1; JOINED.
EMBL; X98315; CAA66653.1; JOINED.
EMBL; X98316; CAA66653.1; JOINED.
EMBL; X98317; CAA66653.1; JOINED.
EMBL; X98318; CAA66653.1; JOINED.
EMBL; X98319; CAA66653.1; JOINED.
EMBL; X98320; CAA66653.1; JOINED.
EMBL; X98321; CAA66653.1; JOINED.
EMBL; X98322; CAA66653.1; JOINED.
EMBL; X98323; CAA66653.1; JOINED.
EMBL; X98324; CAA66653.1; JOINED.
EMBL; X98325; CAA66653.1; JOINED.
EMBL; X98326; CAA66653.1; JOINED.
EMBL; X98327; CAA66653.1; JOINED.
EMBL; X98328; CAA66653.1; JOINED.
EMBL; X98329; CAA66653.1; JOINED.
EMBL; X98330; CAA66653.1; JOINED.
EMBL; X98331; CAA66653.1; JOINED.
EMBL; X98332; CAA66653.1; JOINED.
EMBL; X98333; CAA66653.1; JOINED.
EMBL; X98334; CAA66653.1; JOINED.
EMBL; X98335; CAA66653.1; JOINED.
EMBL; X98336; CAA66653.1; JOINED.
EMBL; X98337; CAA66653.1; JOINED.
EMBL; X98338; CAA66653.1; JOINED.
EMBL; X98339; CAA66653.1; JOINED.
EMBL; X98340; CAA66653.1; JOINED.
EMBL; X98341; CAA66653.1; JOINED.
EMBL; X98342; CAA66653.1; JOINED.
EMBL; X98343; CAA66653.1; JOINED.
EMBL; X98344; CAA66653.1; JOINED.
EMBL; X98345; CAA66653.1; JOINED.
EMBL; X98346; CAA66653.1; JOINED.
EMBL; X98347; CAA66653.1; JOINED.
EMBL; X98348; CAA66653.1; JOINED.
EMBL; X98349; CAA66653.1; JOINED.
EMBL; X98350; CAA66653.1; JOINED.
EMBL; X98351; CAA66653.1; JOINED.
EMBL; X98352; CAA66653.1; JOINED.
EMBL; X98353; CAA66653.1; JOINED.
EMBL; X98354; CAA66653.1; JOINED.
EMBL; X98355; CAA66653.1; JOINED.
EMBL; X98356; CAA66653.1; JOINED.
EMBL; X98357; CAA66653.1; JOINED.
EMBL; X98358; CAA66653.1; JOINED.
EMBL; X98359; CAA66653.1; JOINED.
EMBL; X98360; CAA66653.1; JOINED.
EMBL; X98361; CAA66653.1; JOINED.
EMBL; X98362; CAA66653.1; JOINED.
EMBL; X98363; CAA66653.1; JOINED.
EMBL; X98364; CAA66653.1; JOINED.
EMBL; X98365; CAA66653.1; JOINED.
EMBL; X98366; CAA66653.1; JOINED.
EMBL; X98367; CAA66653.1; JOINED.
EMBL; X98368; CAA66653.1; JOINED.
EMBL; X98369; CAA66653.1; JOINED.
EMBL; X98370; CAA66653.1; JOINED.
EMBL; X98371; CAA66653.1; JOINED.
EMBL; X98372; CAA66653.1; JOINED.
EMBL; X98373; CAA66653.1; JOINED.
EMBL; X98374; CAA66653.1; JOINED.
EMBL; X98375; CAA66653.1; JOINED.
EMBL; X98376; CAA66653.1; JOINED.
EMBL; X98377; CAA66653.1; JOINED.
EMBL; X98378; CAA66653.1; JOINED.
EMBL; X98379; CAA66653.1; JOINED.
EMBL; X98380; CAA66653.1; JOINED.
EMBL; X98381; CAA66653.1; JOINED.
EMBL; X98382; CAA66653.1; JOINED.
EMBL; X98383; CAA66653.1; JOINED.
EMBL; X98384; CAA66653.1; JOINED.
EMBL; X98385; CAA66653.1; JOINED.
EMBL; X98386; CAA66653.1; JOINED.
EMBL; X98387; CAA66653.1; JOINED.
EMBL; X98388; CAA66653.1; JOINED.
EMBL; X98389; CAA66653.1; JOINED.
EMBL; X98390; CAA66653.1; JOINED.
EMBL; X98391; CAA66653.1; JOINED.
EMBL; X98392; CAA66653.1; JOINED.
EMBL; X98393; CAA66653.1; JOINED.
EMBL; X98394; CAA66653.1; JOINED.
EMBL; X98395; CAA66653.1; JOINED.
EMBL; X98396; CAA66653.1; JOINED.
EMBL; X98397; CAA66653.1; JOINED.
EMBL; X98398; CAA66653.1; JOINED.
EMBL; X98399; CAA66653.1; JOINED.
EMBL; X98400; CAA66653.1; JOINED.
EMBL; X98401; CAA66653.1; JOINED.
EMBL; X98402; CAA66653.1; JOINED.
EMBL; X98403; CAA66653.1; JOINED.
EMBL; X98404; CAA66653.1; JOINED.
EMBL; X98405; CAA66653.1; JOINED.
EMBL; X98406; CAA66653.1; JOINED.
EMBL; X98407; CAA66653.1; JOINED.
EMBL; X98408; CAA66653.1; JOINED.
EMBL; X98409; CAA66653.1; JOINED.
EMBL; X98410; CAA66653.1; JOINED.
EMBL; X98411; CAA66653.1; JOINED.
EMBL; X98412; CAA66653.1; JOINED.
EMBL; X98413; CAA66653.1; JOINED.
EMBL; X98414; CAA66653.1; JOINED.
EMBL; X98415; CAA66653.1; JOINED.
EMBL; X98416; CAA66653.1; JOINED.
EMBL; X98417; CAA66653.1; JOINED.
EMBL; X98418; CAA66653.1; JOINED.
EMBL; X98419; CAA66653.1; JOINED.
EMBL; X98420; CAA66653.1; JOINED.
EMBL; X98421; CAA66653.1; JOINED.
EMBL; X98422; CAA66653.1; JOINED.
EMBL; X98423; CAA66653.1; JOINED.
EMBL; X98424; CAA66653.1; JOINED.
EMBL; X98425; CAA66653.1; JOINED.
EMBL; X98426; CAA66653.1; JOINED.
EMBL; X98427; CAA66653.1; JOINED.
EMBL; X98428; CAA66653.1; JOINED.
EMBL; X98429; CAA66653.1; JOINED.
EMBL; X98430; CAA66653.1; JOINED.
EMBL; X98431; CAA66653.1; JOINED.
EMBL; X98432; CAA66653.1; JOINED.
EMBL; X98433; CAA66653.1; JOINED.
EMBL; X98434; CAA66653.1; JOINED.
EMBL; X98435; CAA66653.1; JOINED.
EMBL; X98436; CAA66653.1; JOINED.
EMBL; X98437; CAA66653.1; JOINED.
EMBL; X98438; CAA66653.1; JOINED.
EMBL; X98439; CAA66653.1; JOINED.
EMBL; X98440; CAA66653.1; JOINED.
EMBL; X98441; CAA66653.1; JOINED.
EMBL; X98442; CAA66653.1; JOINED.
EMBL; X98443; CAA66653.1; JOINED.
EMBL; X98444; CAA66653.1; JOINED.
EMBL; X98445; CAA66653.1; JOINED.
EMBL; X98446; CAA66653.1; JOINED.
EMBL; X9844

432 STDNTYMLQINPDSSINPDHLSYFHVGRIMGLAVFHVGHVINGGTFVYKOLIGKPIQL 491
593 D-ESTKLFVFNPSF---ETEGQTLIGVIGLAIYNNCLDVFHFMVYKLMGKGTG 648
492 SDLESVPELHKSILWIL--ENDITPVLDTFCV-EHNAFGRILQHELKPNRNVPVTEE 548
649 RDLGDSHVLVQSLKDLLEVEGVEDDMITFQISQDLFGNPMYDLKENGDKIPITNE 708
549 NKKEVRLVYVNRVFWGIEQFLALQGF---NELIPQHLKPPFQKLELILIGLDKI 604
709 NRKGFVNYSYIILNKSVKQKAFRFGFMVNTSPKLYLFRP---EEILLICGRNL 765
605 DLNDWKSNTLKHCVADSNIVRWFQAVETFDERRARLLQVGTGSTRVPLQGFALQ 663
766 DFOALEETEDYGGVTRDSVLIREFWEIVHSFTDEQKRLFLQTTGTDRAVVGGLKLM 825
664 STGAAGPLFTIHLIDANTDNLKPAHTCENRIDIPPYSEYKLYKLLTAVEETCGFV 722
826 IIAKNGP-----DTERLPTSHTCFNVLLLPYSSKEKLERLLKAITYAKGFGM 874

JUL 11
1A MOUSE
UE3A MOUSE STANDARD; PRT; 885 AA.
008759; P97482;
30-MAY-2000 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Ubiquitin-protein ligase E3A (EC 6.3.2.-) (Oncogenic protein-associated protein E6-AP).
UE3A.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6 X CBA;
MEDLINE=97326076; PubMed=9182527;
Hatakeyama S., Jensen J.P., Weissman A.M.;
"Subcellular localization and ubiquitin-conjugating enzyme (E2) interactions of mammalian HECT family ubiquitin protein ligases.";
J. Biol. Chem. 272:15085-15092 (1997).
[2]
SEQUENCE FROM N.A.
STRAIN=BALB/c; Tissue=Brain;
MEDLINE=97264343; PubMed=9110176;
Sutcliffe J.S., Jiang Y.-H., Galjaard R.-J., Matsuura T., Fang P., Kubota T., Christian S.L., Bressler J., Cattanch B., Ledbetter D.H., Beaudet A.L.;
"The E6-AP ubiquitin-protein ligase (UE3A) gene is localized within a narrowed Angelman syndrome critical region.";
Genome Res. 7:368-377 (1997).
-1- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES.
-1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
-1- TISSUE SPECIFICITY: MOST ABUNDANT IN BRAIN, HEART AND THYMUS.
-1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOESTER FORMATION.
-1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.
-1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-38 IS THE INITIATOR.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U96636; AAB63361.1; --

DR EMBL; U82122; AAB47756.1; --
DR MCD; MGI; 1105098; Ube3a
DR GO; GO:0005829; C:cytosol; IDA.
DR GO; GO:0005834; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding activity; IPI.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
DR InterPro; IPR000569; HECT_domain.
DR Pfam; PF00632; HECT; 1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS00237; HECT; 1.
KW Nuclear protein; Ub1 conjugation pathway; Ligase.
FT DOMAIN 405 410
FT DOMAIN 786 885
FT BINDING 853 853
FT CONFLICT 187 187
FT CONFLICT 301 301
FT CONFLICT 343 343
FT CONFLICT 384 384
FT CONFLICT 460 460
FT CONFLICT 486 486
FT CONFLICT 535 535
FT CONFLICT 577 582
FT CONFLICT 584 584
FT CONFLICT 587 588
FT CONFLICT 597 597
FT CONFLICT 623 627
FT CONFLICT 725 726
FT CONFLICT 817 817
FT CONFLICT 869 870
SQ SEQUENCE 885 AA; 101175 MW; 55D885E090CCB699 CRC64;
Query March 12.4%; Score 483; DB 1; Length 885;
Best Local Similarity 28.5%; Pred. No. 4.4e-27;
Matches 152; Conservative 103; Mismatches 216; Indels 62; Gaps 18;
QY 209 VRGSLQTPQNRPHGHOS-PELPE-----GYEQRTVQGVYFLHTQTGVSTWHDPRIP 260
DB 395 VGGDVDTNHNEDDESPESSELTLQELLGERRNNKGGPRVDPLETELGVKTL-DCRKP 453
QY 261 RDLNVNCDLGLPFGWVRSVSGRIYFVDHNNFTQFTDPRLHHNHOCQLKPSQ 320
DB 454 LISFEESINE--FLNDVLEMDKDYT--FFKVTENKSGSFMTCFFILNAVTKNLGYDNR 509
QY 321 PLPLPSEGLSEDEELPAQRYERDLVKLVRLHSLQLOPQAGHCRIEVSREEIFEEES-- 378
DB 510 -IRMYSE-----RRITVL-YSLVQSQQLAPYLTLYKVRDHIIDALV 549
QY 379 -YRQIMKVRPKDKLMVKFRGEBGLDYGVARWMLYLICHEMLNPNYGLFYQYSTDNIY 437
DB 550 RLEMIAENPADLKKLYVEFEGEQGVSGVSKFQLVWVEEINPDIGMFTYD-EATK 608
QY 438 MLQINPDSSINPDHLSYFHVGRIMGLAVFHVGHVINGGTFVYKOLIGKPIQLSLESV 497
DB 609 LFWNPSF---ETEGQTLIG-ILGLAIYNNCLDVFHFMVYKLMGKGTGPRDLGDS 664
QY 498 DPFLHKSILWIL--ENDITPVLDTFCV-EHNAFGRILQHELKPNRNVPVTEENKKEVY 554
DB 665 HPVLVQSLKDLLEVEGVEDDMITFQISQDLFGNPMYDLKENGDKIPITNEKKEFV 724
QY 555 RLYVNRVFMGIEHQAFLALQGF---NELIPQHLKPPFQKLELILIGLDKIDLNDWK 610
DB 725 ISYDYLNKSVKQKAFRFGFMVNTSPKLYLFRP---EEILLICGRNLDPQALE 781
QY 611 SNTRLK-HCVADSNIVRWFQAVETFDERRARLLQVGTGSTRVPLQGFALQGGTGAAG 669
DB 782 ETTEYDGGYITRESVVIREFWEIVHSFTDEQKRLFLFTTGTDRAPVVGGLKLMIAKNG 841
QY 670 PRLTIHLIDANTDNLKPAHTCENRIDIPPYSEYKLYKLLTAVEETCGFV 722
DB 842 P-----DTERLPTSHTCFNVLLLPYSSKEKLNVLKAITYAKGFGM 884

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D25215; BAA04945.1; --
Genew; HGNC:4876; HERC3.
MIM: 605200; --
InterPro; IPR000569; HECT domain.
InterPro; IPR000408; Reg_Chrcondens.
Pfam; PF00632; HECT; 1.
PRINTS; PR00415; RCC1; 4.
PRINTS; PR00633; RCCNDNSATION.
SMART; SM00119; HECT; 1.
PROSITE; PS0237; HECT; 1.
PROSITE; PS00625; RCC1; 1; FALSE_NEG.
PROSITE; PS00626; RCC1; 2; 4.
PROSITE; PS00012; RCC1; 3; 7.
Ubl conjugation pathway; Ubl conjugation; Repeat.
REPEAT 1 51
REPEAT 2 101
REPEAT 3 102
REPEAT 4 154
REPEAT 5 207
REPEAT 6 208
REPEAT 7 259
REPEAT 8 261
REPEAT 9 311
REPEAT 10 313
REPEAT 11 366
DOMAIN 951
BINDING 1018
SEQUENCE 1050 AA; 117188 MW; 5F08A1DE1F40B912 CRC64;
Query Match 11.6%; Score 449.5; DB 1; Length 1050;
Best Local Similarity 28.5%; Pred. No. 1.5e-24;
Matches 124; Conservative 86; Mismatches 184; Indels 41; Gaps 11;
311 HOCOLK-RPS---QPLPSPSGSLEDELPQRYERDLVOKLVRLHSLQ-----359
633 HQAGKARPSIIQDTVLCSPFFDAQAKTLMQDAEQMVANGLQNVFLLTL 692
360 -----POAGHRIEVSREIFESYRQIMKMPKDLKRLMKVFRGEGLDYGQVAREWLY 415
693 EPLLARSFPLVLRNRLVGDALRELSIHSIDILKKPLKLVFDGEEAVDAGGVTKFEFL 752
416 LICHEMLNPYGLFOYSDNYMLQINPDSSINPDHLSYFHFVGRINGLAFTVHGNGS 475
753 LULKELLNPYGMFTYYQDS-NLWFSDTCEV--EH-NWPHLLGITCGLAYNSTVVDLH 808
476 FTVPPYKQLQKPIQLSDLESVDPELHKLWILE--NDITPVLDTFCVE----HNAP 528
809 FFLALYKLLNVKPGLEDKLSPTGSRSLQELLDYPCEDV---BETFCNFTICRESY 864
529 GRILQHELKPNRNPVTEENKKEVRLYVNRVFRGIEAQFLAQKGFNELIPQHLXP 588
865 GVEQKQLPGDNDVYCKDRQRFVDAYVNVFQISVHEWYTAFFSGFLKVCQGVLEL 924
589 FQKLELIIGLDLIDNDLWKSNTRLK-HCVADSNIVRMFQVAFEDFEERRARLLQFV 647
925 FQPSLRAMVGNVSNWEELEEAIAIKGYSATHPTVKLFWETTFEFPLEKXKLLFL 984
648 TGSTRVLPQGVKALQSGTAGAPRLFTTHLDANTDNLPAKHTCFNRIDIPPVSEYKLY 707
985 TGSDRIPYGMASLQ-----IVIQSTASGEVLPVAHTCYNLLDLPKYSKRELLS 1034
708 EKLLTAVEETCGFAV 722
1035 ARLTQALDNYEGFSL 1049

SULT 14
IB_HUMAN
_TRIB_HUMAN
Q14669; Q15644;
STANDARD; PRT; 1992 AA.

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Thyroid receptor interacting protein 12 (TRIP12).
GN TRIP12 OR KIAA0045.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RN DNA Res. 1:223-229 (1994).
RN [2]
RP SEQUENCE OF 1801-1992 FROM N.A.
RX MEDLINE=95295737; PubMed=7776974;
RA Lee J.W., Choi H.-S., Gyuris J., Brent R., Moore D.D.;
RT "Two classes of proteins dependent on either the presence or absence
RT of thyroid hormone for interaction with the thyroid hormone
RT receptor.";
RL Mol. Endocrinol. 9:243-254 (1995).
CC -!- FUNCTION: COMPONENT OF P4700, AN ATP-DEPENDENT MULTISUBUNIT
CC PROTEIN THAT ACTIVATES THE PROTEOLYTIC ACTIVITIES OF THE
CC MULTIFUNCTIONAL PROTEINASE (20S PROTEASOME) OF THE 26S COMPLEX.
CC SPECIFICALLY INTERACTS WITH THE LIGAND BINDING DOMAIN OF THE
CC THYROID HORMONE RECEPTOR (IN A THYROID HORMONE T3-INDEPENDENT
CC MANNER) AND WITH RETINOID X RECEPTOR (RXR). COULD BE E3 UBIQUITIN-
CC PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-
CC CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND THEN DIRECTLY
CC TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES.
CC -!- SUBUNIT: P4700 IS COMPOSED OF AT LEAST 16 DISTINCT PEPTIDES
CC RANGING IN MOL. WT. FROM 20-112 kDa.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thioester formation.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; D28476; BAA05837.1; --
EMBL; L40383; AAC41731.1; --
DR Genew; HGNC:12306; TRIP12.
DR MIM; 604506; --
DR GO; GO:0016923; F:ligand-dependent thyroid hormone receptor i. . . ; IDA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; NAS.
DR GO; GO:0016567; P:protein ubiquitination; NAS.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR004170; WWE_dom.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF02825; WWE; 1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS00625; RCC1; 1.
DR PROSITE; PS00626; RCC1; 2; 4.
DR PROSITE; PS00012; RCC1; 3; 7.
DR Ubl conjugation pathway; Ligase.
KW Proteasome; Ubl conjugation pathway; Ligase.
FT DOMAIN 1885 1992
FT BINDING 1959 1959
FT BINDING 1959 1959
FT CONFLICT 1969 1992
FT CONFLICT 1969 1992
SQ SEQUENCE 1992 AA; 220432 MW; 294A7C063A332DE CRC64;
Query Match 9.2%; Score 358; DB 1; Length 1992;
Best Local Similarity 25.1%; Pred. No. 1.7e-17;

arch completed: February 20, 2004, 15:26:05
o time : 15.8131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 protein - protein search, using sw model
in on: February 20, 2004, 15:25:26 ; Search time 33.9137 Seconds
(without alignments)
5501.382 Million cell updates/sec

tle: US-10-009-945-2
irect score: 3884
quence: 1 GGSIKIRLVLCANLAKK.....EKLYEKLTAVBETCGFAVE 723

oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 830525 seqs, 258052604 residues

tal number of hits satisfying chosen parameters: 830525

imum DB seq length: 0
imum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase : SPTREMBL 23:**
1: sp_arChaea:*
2: sp_Bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_Bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3792.5	97.6	728	11	Q8K300 mus musculus
2	2206	56.8	1061	5	Q9V853 drosophila
3	1895	48.8	355	11	Q8BSC0 mus musculus
4	1559.5	40.2	295	4	Q8NDD8 homo sapien
5	1341	34.5	786	3	O14326 schizosach
6	1340.5	34.5	288	4	Q96DE7 homo sapien
7	1318	33.9	854	11	O54971 mus musculus
8	1301	33.5	862	4	Q9BY75 homo sapien
9	1288.5	33.2	903	4	Q96F66 homo sapien
10	1269	32.7	911	4	Q8NSA7 homo sapien
11	1247	32.1	955	4	Q96F05 homo sapien
12	1245.5	32.1	703	4	Q9H451 homo sapien
13	1245.5	32.1	739	4	O43584 homo sapien
14	1241.5	32.0	820	4	Q9NT88 homo sapien
15	1239	31.9	922	4	Q9H0M0 homo sapien
16	1239	31.9	949	5	Q9Y0H4 drosophila

17	1234	31.8	870	11	Q9DBH0 mus musculus
18	1228.5	31.6	855	4	Q8WU09 homo sapien
19	1228.5	31.6	858	4	Q9BW58 homo sapien
20	1228.5	31.6	971	13	O42573 xenopus lae
21	1228.5	31.6	995	4	O43165 homo sapien
22	1227.5	31.6	887	11	Q62940 rattus norv
23	1226	31.6	854	4	Q9H2M4 homo sapien
24	1225	31.5	870	4	Q96CZ2 homo sapien
25	1225	31.5	870	4	O00308 homo sapien
26	1225	31.5	887	11	Q8BGB3 mus musculus
27	1224	31.5	855	11	Q8CF10 mus musculus
28	1223.5	31.5	835	11	Q8BRT9 mus musculus
29	1215	31.3	855	11	Q99PK2 mus musculus
30	1214.5	31.3	834	5	Q9SR64 drosophila
31	1213	31.2	794	5	Q9N2Z7 caenorhabdi
32	1210	31.2	792	5	Q9SXU3 caenorhabdi
33	1207.5	31.1	376	13	Q98T84 xenopus lae
34	1200.5	30.9	258	11	Q9CSE3 mus musculus
35	1196.5	30.8	617	11	Q8BZ23 mus musculus
36	1196.5	30.8	956	5	Q8TOR6 drosophila
37	1195.5	30.8	518	5	Q8T0C8 drosophila
38	1191	30.7	1007	5	Q9VVI3 drosophila
39	1189	30.6	777	11	O08758 mus musculus
40	1151	29.6	838	5	Q9STQ0 drosophila
41	1129	29.1	724	5	Q9BKW4 caenorhabdi
42	1047	27.0	684	4	O00307 homo sapien
43	1026.5	26.4	671	3	Q9UTG2 schizosach
44	1002	25.8	1585	4	Q9HCC7 homo sapien
45	991	25.5	759	11	Q9C863 mus musculus

ALIGNMENTS

RESULT 1

Q8K300 PRELIMINARY; PRT; 728 AA.

AC Q8K300; (T-EMBLrel. 22, Created)

DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)

DE Similar to E3 ubiquitin ligase SMURF1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

[1]

RP SEQUENCE FROM N.A.

RA Strausberg R.

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.

DR EMBL; BC029097; AAH29097.1; -.

DR InterPro; IPR000008; C2.

DR InterPro; IPR000569; HECT domain.

DR InterPro; IPR001202; WW_Rsp5_WWP.

DR Pfam; PF00168; C2.1.

DR Pfam; PF00632; HECT; 1.

DR Pfam; PF00397; WW; 2.

DR SMART; SM00239; C2; 1.

DR SMART; SM00119; HECTC; 1.

DR SMART; SM00456; WW; 2.

DR PROSITE; PS00499; C2 DOMAIN_1; 1.

DR PROSITE; PS00004; C2_DOMAIN_2; 1.

DR PROSITE; PS0237; HECT; 1.

DR PROSITE; PS01159; WW DOMAIN_1; 1.

DR PROSITE; PS0020; WW_DOMAIN_2; 2.

DR Ligase.

SQ SEQUENCE 728 AA; 83083 MW; C355291B9D8AD757 CRC64;

Query Match 97.6%; Score 3792.5; DB 11; Length 728;
Best Local Similarity 98.1%; Pred. No. 2e-303;
Matches 708; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

2 GSSIKILTLVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTDTKNTKLTDPKWNHXYDLY 61
10 GSSIKILTLVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTDTKNTKLTDPKWNHXYDLY 69
62 VGKTDSTISVNNHKKHKGAGFLGCVRLLSNAISRLKOTGYORLDCCKLPNSDTPAV 121
70 VGKTDSTISVNNHKKHKGAGFLGCVRLLSNAISRLKOTGYORLDCCKLPNSDTPAV 129
122 RGQIVSLQTRDRLGTGSGVVDRCRGLLENGTVYEDSGRPLSCFMEPEAPYDSTGAA 181
130 RGQIVSLQTRDRLGTGSGVVDRCRGLLENGTVYEDSGRPLSCFMEPEAPYDSTGAA 189
182 AGGNCRFVSPSQDORLQALRNPVDRVSLQTPQNRPHGQSPPELPGYEQRTTVQGG 241
190 AGGNCRFVSPSQDORLQALRNPVDRVSLQTPQNRPHGQSPPELPGYEQRTTVQGG 249
242 VFLLHTGTGTVSTHDPRIPLDLSVNCDELGPPLPGWEVSTVSGRIYFVDNNRTTFT 301
250 VFLLHTGTGTVSTHDPRIPLDLSVNCDELGPPLPGWEVSTVSGRIYFVDNNRTTFT 309
302 DPLRLHIMNHQCOKEPSQPLPSPSEGSLEDELPQRYERDLVQKLKVLRLHLSLQOPQ 361
310 DPLRLHIMNHQCOKEPSQPLPSPSEGSLEDELPQRYERDLVQKLKVLRLHLSLQOPQ 369
362 AGHCRISVREETFEESYRQIMKWRPKDLKRLMWKFRGEBGLDTGGVAREWLYLLCHEM 421
370 AGHCRISVREETFEESYRQIMKWRPKDLKRLMWKFRGEBGLDTGGVAREWLYLLCHEM 429
422 LNPYVGFQFQSTONIVMLQNPDSINPHLSYFHFVGRIMGLAVFHGYINGFTVPFY 481
430 LNPYVGFQFQSTONIVMLQNPDSINPHLSYFHFVGRIMGLAVFHGYINGFTVPFY 489
482 KOLLGKPIQLSDIESVDPELHKLWILENDITPVLVDHTFCVEHNAFRILQHELPKNGR 541
490 KOLLGKPIQLSDIESVDPELHKLWILENDITPVLVDHTFCVEHNAFRILQHELPKNGR 549
542 NVPVTENKKEYVRLVYNNRFGMIGIAQFALQKGFNELIPQHLKPKDQKELELILGGL 601
550 NVPVTENKKEYVRLVYNNRFGMIGIAQFALQKGFNELIPQHLKPKDQKELELILGGL 609
602 DKIDLNDKNTLKHCVADSNIVRVFWQAVETFDERRARLLQFVTGSTRVLPQGFAL 661
610 DKIDLNDKNTLKHCVADSNIVRVFWQAVETFDERRARLLQFVTGSTRVLPQGFAL 669
662 QGSGTAGPRLFTIHLIDANTDNLKPAHCFNRIDIPPYESYEKLYEKLITAVEETCGFA 721
670 Q---GAAGPRLFTIHLIDANTDNLKPAHCFNRIDIPPYESYEKLYEKLITAVEETCGFA 726
722 VE 723
727 VE 728

SULT 2

V853

PRELIMINARY;

PRT; 1061 AA.

Q9V853;

Q9V853; Q9V3W2;

01-MAY-2000 (T-EMBLrel. 13, Created)

01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

01-MAR-2003 (T-EMBLrel. 23, Last annotation update)

CG4943 Protein (E3 ubiquitin ligase) (ubiquitin-protein ligase) (Smad-

ubiquitin E3 ligase Smurf1).

LACK OR SMURF OR SMURF1 OR CG4943.

Drosophila melanogaster (Fruit fly).

Eukaryota; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Echydnoidea; Drosophilidae; Drosophila.

NCBI_TaxID=7227;

[2]

SEQUENCE FROM N.A.

STRAIN=Berkeley;

MEDLINE=20196006; PubMed=107311132;

Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Brill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burdick K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong P., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wessman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng J.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupu J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Skarlat S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

```
[6]
1 SEQUENCE FROM N.A.
2 Laurence A., Hawley S.;
3 "Molecular cloning of a type E3 Ubiquitin ligase.";
4 Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
5 [7]
6 SEQUENCE FROM N.A.
7 Podos S.D., Hanson K.K., Wang Y.-C., Ferguson E.L.;
8 "The DSmurf ubiquitin-protein ligase restricts BMP signaling spatially
9 and temporally during Drosophila development.";
10 Dev. Cell 1:0-0 (2001).
11 [8]
12 SEQUENCE FROM N.A.
13 Liang Y.-Y., Lin X., Feng X.-H.;
14 "dSmurf, a Smad-ubiquitin E3 ligase, specifically targets dpp-
15 activated Mad protein for degradation.";
16 Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
17 -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
18 EMBL; AB003802; AAF57824.3; -
19 EMBL; AF216521; AAF21125.1; -
20 EMBL; AF416571; AAL09691.1; -
21 EMBL; AF464851; AAMC9646.1; -
22 HSSP; Q13526; 12IN.
23 Flybase; FBgn029006; lack.
24 InterPro; IPR000008; C2.
25 InterPro; IPR000569; HECT_domain.
26 InterPro; IPR002349; WW.
27 InterPro; IPR001202; WW_Reps_WWP.
28 Pfam; PF00169; C2; 1.
29 Pfam; PF00397; WW; 3.
30 PRINTS; PR00403; WWDOMAIN.
31 SMART; SM00239; C2; 1.
32 SMART; SM00456; WW; 3.
33 PROSITE; PS00499; C2_DOMAIN_1; 1.
34 PROSITE; PS00004; C2_DOMAIN_2; 1.
35 PROSITE; PS00237; HECT; 1.
36 PROSITE; PS01159; WW_DOMAIN_1; 2.
37 PROSITE; PS00020; WW_DOMAIN_2; 3.
38 Ligase.
39 SEQUENCE 1061 AA; 115675 MW; 68BCC550F5129163 CRC64;
40
41 Query Match 56.8%; Score 2206; DB 5; Length 1061;
42 Best Local Similarity 43.0%; Pred. No. 1.8e-172;
43 Matches 454; Conservative 108; Mismatches 148; Indels 346; Gaps 14;
44
45 6 KIRTVLCANLAKQFFRLPDPFAKIVVDGSGQCHSTDTVKNLDPKKNQHYDIYVGT 65
46 14 KVRITILCAENLARKDLFLPDPFAKIVVDGSGQCHSTDTVKNLDPKKNQHYDIYVGT 73
47
48 66 DSITISVNNHKKTHKKGAGFLGCVRLLSNAISRLKDTGYORLDLCKLNPSTDAVRQI 125
49 74 DAITITVNNRKLHK--GSGFLGCVRIAPAFNQLSKGAGFQRLDGLKSPDDDLVRQI 131
50
51 126 VSLQTRDRIGTG-----GSVVDCRGLLENE-----GTVY----- 155
52 132 IISLLSKDGPSSGNPLAIVGSDVRGSDSDSDSLPEGWEERTDNGRYVYVNHATK 191
53 156 ----- 155
54
55 192 STQMDRPRQPGVGVSSHATSPQORHNTHNGSDRQAPAGTRSTCTNLMMNHRSDL 251
56 156 -----EDSGKRLPSCFMEEPAPYTDSTGAAAGGNCRFVE----- 191
57
58 252 SVTASDERRHSTILSVGKENTSPPTVSA--TTPGKKTSSNSAGG--RTLEQRT 308
59
60 192 ----SPQDRLQARLNRNDVKGSLQTPQNRPHGQSP----- 227
61 309 NEPATPSSTTSASVRLHNSD--NHVKTPKQTNCHAPPESTPTSGQNVYVNGAQN 366
62 228 ----- 227
63
64 367 STSGNGSGQAQPCASANGTQDAATTTSPSTTSPPHSHQSPPTPNISPPASVTPSAN 426
```

```
228 ----- 227
427 GNVHSPNANSTPAGSGGSRSYTAATPQSRQRRSSROQGESESTRRRSSRGTRNGTSG 486
228 -----LPGYEORTTVQGVYFLHTQTGVSTWHPRI 259
487 GGGGGSGQRYASAAIAAQAARPLDLPGYEWRTTQGGVYFYHPTGVSTWHPRI 546
260 PDLNS--VNCDELGPLPGWEVSTVSGRIYFVDHNNRTTQFTDPLHHIM----- 309
547 PRDFDTQHLTLDAIGLPSGWEQRKTASGRVYFVDHNNRTTQFTDPLSGSILQWIRRG 606
310 ---NHQCLKEPSQPLPLPS-----EG 328
607 VPTTSANAGTAPPSATPAPSAATAVPQATPASNATPTTLTTNPPHRIYVDLPQG 666
329 SLEDELPQARYERDLVQKLKVLRLHLSLQOPQAGHCHREIVSREEIFEESYRQIMKRPK 388
667 LLEGADL-LPKYRRDLVGKRALRTELQTMQPSGHCHREIVSREEIFEESYRQIMKRAK 725
389 DLKRLMWKFERGEGLDYGVAREWLYLLCHEMLNPPYGLFOYSTDNIMYMLQINPDSSIN 448
726 DMKRLMWKFERGEGLDYGVAREWLYLLCHEMLNPPYGLFOYSTDNIMYMLQINPDSSIN 785
449 PDHLSYFHFVGRIMGLAVFHGHYINGGTFVPFYKQLGKPIQLSDLESVDPELHKSIAWI 508
786 PDHLSYFHFVGRITGLIAVFGHCLDGGFTTFFYKQLGKPIQLSDLESVDPELHKSIAWI 845
509 LENDITPVLDTFCVEHNAFGRILQHELKPNGRNVPTTEENKKEYVRLYNWRFMRGIEA 568
846 LESNISGIIESTFSVENNSFGALVHVELKPGGASIPVTEENKKEYVRLYNWRFMRGIEA 905
569 QFLALQKGFNELIPOHLLKPPFOKELELIIGGLDKIDLNDWKSNTLKHCVADSNIVRWF 628
905 QFLALQKGFCELIPLSHLLRPDERELELVIGGISIDVNDWRNTRLKHCVADSNIVRWF 965
629 WQAVETFDERRARLLQFVTGSTRVPLQKALQSGTGAAGRLFTIHL-IDANTDMLPK 687
966 WQVYESYSSEMARLLQFVTGSTRVPLQKALQSGTGAAGRLFTIHL-IDANTDMLPK 1025
688 AHTCFNRIDIPPYSEYKELYKLLTAVEETCGFAVE 723
1026 AHTCFNRIDIPPYETIYQLLCKLTOAVEETCGFAVE 1061
RESULT 3
QBSC0 PRELIMINARY; PRT; 355 AA.
AC QBSC0,
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical HECT domain (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium.
RA THE RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK034736; BAC28813.1; -.
KW Hypothetical protein.
FT NON_TER
FT 1
SQ SEQUENCE 355 AA; 41315 MW; 93B39B4C82F86DD1 CRC64;
Query Match 48.8%; Score 1995; DB 11; Length 355;
Best Local Similarity 99.7%; Pred. No. 1.4e-147;
```

Matches 354; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

369 VSREBIFESYQIMKMPKOLKKRLMWKFRGEGLDYGGVAREWLYLLCHEMLNPFYGL 428
1 VSREBIFESYQIMKMPKOLKKRLMWKFRGEGLDYGGVAREWLYLLCHEMLNPFYGL 60

429 FOYSTDNIYMLQINPDSSINPDHLSYFHFVGRINGLAVFHHGYINGGFTVPFYKQLLGP 488
61 FOYSTDNIYMLQINPDSSINPDHLSYFHFVGRINGLAVFHHGYINGGFTVPFYKQLLGP 120

489 IQLSDLESVDPELHLSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPNGRNVPTEE 548
121 IQLSDLESVDPELHLSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPNGRNVPTEE 180

549 NKKEYVRLYVNRVFRGIEAQFALQKGFNELIPQHLKPPDKQKELIIGGLDKIDLND 608
181 NKKEYVRLYVNRVFRGIEAQFALQKGFNELIPQHLKPPDKQKELIIGGLDKIDLND 240

609 WKSNTRLKHCVDASNIVRMFQAVETFEERRARLLQFVTGSTRVPLQGFALQGGSTGAA 668
241 WKSNTRLKHCVDASNIVRMFQAVETFEERRARLLQFVTGSTRVPLQGFALQGGSTGAA 300

669 GRLFTIHLIDANTDNLPKAHTCFNRIDIPPYSEYKLYEKLITAVEETCGFAVE 723
301 GRLFTIHLIDANTDNLPKAHTCFNRIDIPPYSEYKLYEKLITAVEETCGFAVE 355

SUULT 4
NDD8 QNDD8 PRELIMINARY; PRT; 295 AA.

QNDD8;
01-OCT-2002 (TrEMBLrel. 22, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Hypothetical protein (Fragment).
DUF2564H223.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Brain;
Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AL834242; CAD38919.1; --
InterPro; IPR000569; HECT_domain.
Pfam; PF00632; HECT; 1.
SMART; SM00119; HECTC; 1.
PROSITE; PS50237; HECT; 1.
Hypothetical protein.
NON_TER
1
SEQUENCE 295 AA; 34030 MW; 0507325127A943EA CRC64;

Query Match 40.2%; Score 1559.5; DB 4; Length 295;
Best Local Similarity 98.7%; Pred. No. 4.6e-120;
Matches 293; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

427 GLFOYSTDNIYMLQINPDSSINPDHLSYFHFVGRINGLAVFHHGYINGGFTVPFYKQLL 486
2 GLFOYSTDNIYMLQINPDSSINPDHLSYFHFVGRINGLAVFHHGYINGGFTVPFYKQLL 61

487 KPIQLSDLESVDPELHLSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPNGRNVPVT 546
62 KPIQLSDLESVDPELHLSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPNGRNVPVT 121

547 EENKXEYRLYVNRVFRGIEAQFALQKGFNELIPQHLKPPDKQKELIIGGLDKIDL 606
122 EENKXEYRLYVNRVFRGIEAQFALQKGFNELIPQHLKPPDKQKELIIGGLDKIDL 181

607 NDWKSNTRLKHCVDASNIVRMFQAVETFEERRARLLQFVTGSTRVPLQGFALQGGSTG 666
182 NDWKSNTRLKHCVDASNIVRMFQAVETFEERRARLLQFVTGSTRVPLQGFALQ--G 238

QY 667 AAGRLFTIHLIDANTDNLPKAHTCFNRIDIPPYSEYKLYEKLITAVEETCGFAVE 723
DB 239 AAGRLFTIHLIDANTDNLPKAHTCFNRIDIPPYSEYKLYEKLITAVEETCGFAVE 295

RESULT 5
O14326 PRELIMINARY; PRT; 786 AA.
AC O14326;
DT 01-JAN-1999 (TrEMBLrel. 09, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ubiquitin-protein ligase.
GN SPBC1659.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Volckaert G., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; 299759; CAB16903.1; --
DR HSP; Q1326; 1P1N
DR GeneDB_Spombe; SPBC1659.11C; --
DR InterPro; IPR000008; C2.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PRO0403; WMDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS50020; WW_DOMAIN_2; 3.
KW Ligase.
SQ SEQUENCE 786 AA; 89259 MW; 57B7A859F5497B9A CRC64;

Query Match 34.5%; Score 1341; DB 3; Length 786;
Best Local Similarity 36.6%; Pred. No. 2.1e-101;
Matches 304; Conservative 128; Mismatches 236; Indels 162; Gaps 21;

QY 6 KIRLTVLCANLAKKDFRPPFAKIVVDGSGQCHSTDTVNTLDPKWNQHYDLVYKGT 65
DB 7 RVRFYVAADGLSKRDLFRQDPFAILTVDGE-QTHTTKVIKSVNPFYVNEGEVTVKPS 65

QY 66 DSITTSVWNHKKHKKQAGFLGCVRLLSNAISRLKDTGYQR-----LDLCK-LNPSDID- 119
DB 66 SVIS-RLFDQKKF-KKKDQGLGLVVF-----RMREVGSFRSNREVSLRPLKKSSTN 118

QY 120 -AVRQIYV-----SL 129
DB 119 LSVLGNLVKVPKIRAPAGNHSSTTANRTTPTTTTARTTTRTPRTATTNTSNQST 178

QY 130 QTRDRIGTGGVDCRGLENHGVYEDSGFRPL-----SCWEEPA-----PYTDS 177
DB 179 SNSTNGTSAATNGTGTGAGTGASHRSS-----PVTNRQTNNTSALSNNAHIMSFEQ 234

QY 178 TGAAGG-----GNCRFVE-----SPSQDQRLQAORLNPV 209
DB 235 YGRLPFGVERRADSLGRYYVDHNTRTTTWTTPASSTNPVHTSSDSQRLNHLNLPD- 293

QY 210 RGSLOTQPNRPHGHQSP---ELPEGYQRTVQGVYFLHTGTGVSTWHDPRIP-----260

```
294 -----DSNPSLMQSDSGNDLFPGEWRYTDTGRPYFVDHNRITTTWDRPNLVRNG 346
261 -----RDLSNVCDELGPPLPGWYRSTVSGRIYFVDHNRITTTQFTDPRHLHIMHOCQ 314
347 GSSTVGLMOPQSLHGLPLSPGWEMRLNSARVYFVDHNTKTTTWDPRL----- 397
315 LKPSQPLPLPSGSLDEELPAQRVERDLVQKLVLRHLSLQOPOAGHCEIIVSREI 374
398 -----PSALDQVP--QYKDFRRLKIYFRSQPM-RPLPQCNVNRVRRDHI 441
375 FEBSYQIMQRPDKLKLMLWKFGEGLDGGVAREWLYLLCHEMLNPNYGLFOYSTD 434
442 FEDSYABIMKYSADLKLMLIRFDEGDLGYGLSREFFLSHKMPDPIYCLFEYSAV 501
435 NIYMLQINPDSSINPHLSYFHFVGRIMGLAVFHHYINGGTFVPFYKQLLQKIQISDL 494
502 DNTLQINPHSSINPHLANYFRIGRVLGIAIIFHRFLDAFFVWSLYKKLIRKYSVSLDM 561
495 ESVDPELHKSILVWLENDITPVLDDHTFCVHNFAFRILQHELKPNRNVVPTENKKEYV 554
562 ESIDABFYSRLKWLENDITGILDLTFSVEEDHFGVTRVELITNGENIETVENKKEYV 621
555 RLYNVRFRGIEAQLALQKGFNELIPOHLKPFQKLELIIGGLDKIDLNDWKSNT 614
622 DLYTEVRVSKRVQPNFYSGFVELVSPDLVNFDERELLLIGISDVVDVWKSHT 681
615 LKHCVADSNIVRFWQAVTFDEERRARLLQFVTGSTRVPLQGFKALQSGAAGPRL 674
682 YRYIATDPVKKFWELIAGWKNRDSKLLQATGTSRIPVNGFRDLQSD--GPRKFT 738
675 IHLIDANT-DNLPKANTCNFRIDIPYSEYKLYEKLITAVEETCGFAVE 723
739 IE--KAGTDPQLPVAHTCNFRIDLPYPSKDLTLEKLSLAVENTVGFNGE 786

SULT 6
6DE7
Q36DE7 PRELIMINARY; PRT; 288 AA.
Q36DE7;
01-DEC-2001 (T-EMBLrel. 19, Created)
01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
Similar to E3 ubiquitin ligase SMURF2 (Fragment).
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Lung;
Strausberg R.;
Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; BC009527; AAH09527.1; -.
InterPro; IPR000569; HECT_domain.
Pfam; PF00632; HECT; 1.
SMART; SM00119; HECTC; 1.
PROSITE; PS50237; HECT; 1.
Ligase.
NON_TER 1
SEQUENCE 288 AA; 33255 MW; FE2B43E300DE66537 CRC64;
Query Match 34.5%; Score 1340.5; DB 4; Length 288;
Best Local Similarity 84.8%; Pred. No. 4.9e-102;
Matches 246; Conservative 23; Mismatches 18; Indels 3; Gaps 1;
```

```
QY 554 VRLYVNRFRMGIEAQLALQKGFNELIPOHLKPFQKLELIIGGLDKIDLNDWKSNT 613
DB 122 VRLYVNRFRMGIEAQLALQKGFNEVIPQHLKTFDEKELELIICGLKIDVNDWKVNT 181
QY 614 RUKHCVADSNIVRFWQAVTFDEERRARLLQFVTGSTRVPLQGFKALQSGAAGPRL 673
DB 182 RUKHCVADSNIVRFWQAVTFDEERRARLLQFVTGSTRVPLQGFKALQSGAAGPRL 238
QY 674 TIHLIDANT-DNLPKANTCNFRIDIPYSEYKLYEKLITAVEETCGFAVE 723
DB 239 TIHLIDANT-DNLPKANTCNFRIDIPYSEYKLYEKLITAVEETCGFAVE 288

RESULT 7
CS4971 PRELIMINARY; PRT; 854 AA.
ID 054971;
AC 054971;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Ubiquitin protein ligase.
GN ITCH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=C3H/HeJ; TISSUE=Kidney;
MEDLINE=98122574; PubMed=9462742;
Perry W.L., Hustad C.M., Swing D.A., O'Sullivan T.N., Jenkins N.A.,
Copeland N.G.;
"the itchy locus encodes a novel ubiquitin protein ligase that is
disrupted in al8H mice";
RL Nat. Genet. 18:143-146(1998).
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AF037454; AAB99764.1; -.
DR HSSP; Q13526; 1P1N.
DR MGD; MGI:1202301; Itch.
DR InterPro; IPR000008; C2.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
KW Ligase.
SQ SEQUENCE 854 AA; 97975 MW; AA832D7BF0B9B3D8 CRC64;
```

```
Query Match 33.9%; Score 1318; DB 11; Length 854;
Best Local Similarity 35.4%; Pred. No. 1.9e-99;
Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;

QY 6 KIRLVLCAL-NLAKOFFRPLDPFAKIVVVGSGQCHSTDTVKNLTDPKNOHYDLYVGK 64
DB 9 QICITIVISAKUKENKWF-G-PSPIVEVTVD--GQSKKTEKKNNTSPRWKQPLTIVTP 65
QY 65 TDSITISVYNNHKIKKQAGFLGCVRL-----LSNAISRLKDTGYQR----- 107
DB 66 TSKLCFRVWSHQL---KSDVLLGTAGLDIVETLKSNNMKLEEVWVTLQVGGKEPTM 122
QY 108 ---LDLCKLNPSDPAVRGQIVS-----LQTRD--RIGTCGS---VVDGR 145
DB 123 GDLSYC-----LDGLQVEAEVVTNGTSCSESTTQNDGCRTRDTRVSTNGSDPEVAAS 178
```

146 G-----LLENEGTVYDSGPGPLSCFMEERAPYT-----DSTG 179
179 GENKRANGNNPSLNGG--FKPSRPPRP-----SRPPPTPRPASVNGSPSTNSDGG 231
180 AA-----AGGNCRFVESPSQ-----DORLAQR 203
232 SSTGSLPPTNTVNTSTSGATGIIILITSGGSPRLNTVQALPPQNEGRV-----287
204 LRNPDPVRSGL-----QTPQNRPHGQSPPELPEGYEQRTTVQGVYFLHTQTGVSTWH 255
288 -----DQGRVYVVDHVEKRTTWDRE-----PLPGWERRVDNNGRIYVVDHFTRTTTWQ 338
256 DRIP-----RDJ-----NSVNCDELGLPPGKEVR 281
339 RPTLESVRVYEQWLORSLOQAMQFQNRFYNGQDLFATSONKEFDPDPLGLPPGWEKR 338
282 STVSGRIYFVDHNRRTTQTPRLHHIMHQQLKEPSQIPLP-----324
399 TDSNGRVYFVNHTRIQWEDR-----SQQLNE--KPLPEGWEMRFTVDGIPYFVDH 450
325 -----PSEGSLEDELPQRYERDLVQKLVLR---HELSQLQOAGHCHIEVSRE 372
451 NRATTYIDERTGKSALDNGPQIAYRDFKAKVQYFRFWCQQLAMPQ---HIKITVTRK 506
373 EIPESVQIOMKRPDKLKRIMVFRGEEGLDYGAVAREWLYLLCHEMLNPPYGLFOYS 432
507 TLPEDSFQIMSFSPDLRRRLVWIPPEEGLDYGAVAREWFFLLSHEVLNPMYCLPEYA 566
433 TDNIYMLQNPSSINPDHLSVTFHVRIMGLAVPHGYINGFTVPYKOLLGKPIOLS 492
567 GKDNVCLQINPASYINPDHLKYFRFTRGFANALPHGKFIDTGFSLPFYKRLNKPVLGK 626
493 DLBSVDPELHKSILWLENDITPV-LDHTFCVEHNAFGRILQHELKPNGRNVPYTEKX 551
627 DLESIDPEFINSILWKENNIECGLEMYFSVDKEILGEIKSHDKPNNGNLTVEBKE 686
552 EYVRLVYVNRFRGIBAQFLAQGFNELIPQHLKLPDQKELELIIGGLDKIDLNDWKS 611
687 EYIRVMAEWRLSRGVEQQAFFEGFNEILPQYLYQYFDAKELEVLCCQMEIDLNDWQR 746
612 NTELKCVADSNIVRFQAVTTEDEERARLLQFVGTSTVPLQGEKALOGSGAGR 671
747 HAIYRHYTTSKQIMFWQFVKEIDNEXRMRLQVFTGTCRLPVGGFADLMGNS---GPQ 803
672 LFTIHLIDANTNLPKXHTCFNRIDIPPEYSEKLYEKLITAVBETCGFAVE 723
804 KFCIEKV-GKENWLPSSHCFNRDLDPYKSYEQLEKLLFAIETETEGFGOE 854
Q9BY75 PRELIMINARY; PRT; 862 AA.
Q9BY75; 01-JUN-2001 (T-EMBLrel. 17, Created)
01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
Ubiquitin protein ligase Itch.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
Miyazaki K., Okamoto Y., Sakamoto M., Nakagawara A.;
"Homo sapiens mRNA for ubiquitin protein ligase Itch, complete cds.";
Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
MEDLINE=20549573; PubMed=10940313;
Qiu L., Joazeiro C., Fang N., Wang H.Y., Elly C., Altman Y., Fang D.,
Hunter T., Liu Y.C.;

RT RT Recognition and ubiquitination of Notch by Itch, a hec-type E3
RL RT ubiquitin ligase.";
RN J. Biol. Chem. 275:35734-35737(2000).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98122574; PubMed=9462742;
RA Perry W.L., Hustad C.M., Swing D.A., O'Sullivan T.N., Jenkins N.A.,
RA Copeland N.G.;
RT "The Itchy locus encodes a novel ubiquitin protein ligase that is
RT disrupted in a18H mice.";
RL Nat. Genet. 18:143-146(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=20501262; PubMed=11046148;
RA Winberg G., Matskova L., Chen F., Plant P., Rotin D., Gish G.,
RA Ingham K., Erberig L., Pawson T.;
RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
RT protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases.";
RL Mol. Cell. Biol. 20:8526-8535(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=98313405; PubMed=9647693;
RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
RA Kaminsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.;
RT "Atrophin-1, the DRELA gene product, interacts with two families of WW
RT domain-containing proteins.";
RL Mol. Cell. Neurosci. 11:149-160(1998).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=21219930; PubMed=11318644;
RA Chen X., Wen S.C., Fukuda M.N., Gavva N.R., Hsu D.W., Akama T.O.,
RA Yang-Peng T.L., Shen C.K.J.;
RT "Human ITCH is a Co-Regulator of the Hematopoietic Transcription
RT Factor NF-E2.";
RL Genomics 73:238-241(2001).
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AB056663; BAB39389.1; -.
DR EMBL; AF095745; AAK39399.1; -.
DR HSSP; Q13526; IPIN.
DR InterPro; IPR000008; C2.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW_domain.
DR InterPro; IPR001202; WW_Rap5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
KW Ligase.
SQ SEQUENCE 862 AA; 98675 MW; A3D96CE7F4DBF9D3 CRC64;
Query Match 33.5%; Score 1301; DB 4; Length 862;
Best Local Similarity 34.7%; Pred. No. 4.8e-98;
Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;
QY 6 KIRLTVLCAK-NLAKDFFRLPDPFAKIVVDGSGQCHSTDTVNTLDPKWNQHYDLYVGK 64
DB 19 QLOITVISAKLKENKNWFG-PSYVEVTD--GQSKTEKCNNTNSPKWQPLTVITVP 75
QY 65 TDSITISVNNKTHKK--QGAGFLGCVRLLSNAISRLKDT-----GYQR-----L 108
DB 76 VSKLHFRVWSHQTLKSDVLLGTAALDIYETLKSNNMKLEEVVTVLQGGDKPTETIGDL 135
QY 109 DLCKLNPSDITDAVRGQIVWSLOT-----RDRIGTGGG-----VWDC 144
DB 136 SIC-----LDGLQLESEVVVTNGETTCSESASQNDGGRSKDETRVSTNGSDPDADAGEN 191


```
145 RGL-----LENGTVEYEDSGRPLSCMERBAPYTDSTGAAGGNCRFVES-----192
192 RRVGNNSPSLNCG--FKSRPRP-----SRPPPTPRPASVNGSPSATSSESDGSGST 244
193 -----
245 GSLPPTNTNTSEGAATGLIIPLTISGGSPRLNPVTQAPLPPGWEQV-----PSQDRLQAQLRNP 208
209 VRGSL-----QTPONRPHGHSPELPGYQRTTVOGVYFLHTQTGVSTWHDPRIP 260
297 QHGRVYVVDHVEKTTWDRPE-----PLPPGWERVDNMGRIIYVDHFTRTTWTQRPFL 351
261 -----RDL-----NSVNCDELGPLPPGWEVASTVSG 286
352 SVRNYEQWLQORSOLQAGMOQFNORFYIGNODLFATSSQKEFDPLGPLPGWEKRTDSNG 411
287 RIYVDHNNRTTQDRLHHIMHQCQLKEPSQPLPL-----SOGQLE--KPLPEGWEMRFTVDGIFYFVDHNNRTT 463
412 RYVFNENTRITQWEDR-----SOGQLE--KPLPEGWEMRFTVDGIFYFVDHNNRTT 463
325 ----PSGSELEDELPQAQYERDLVOKLVLR---HELSQLQPOAGHCHRIEVSREEIFEE 377
464 TYIDPRTGKSALDNGPOIAYVDFKAKVQYFRFWCQOLAMPQ---HIKITVTRKTLFED 519
378 SYROIWKRPDLKRLMVKFRGEBGLDYGVAEWLYLLCHEMLNPPYGLFQYSTDNII 437
520 SFQIMGFSQDRLRRRLWVFPGEGLDYGVAEWFLLSHEVLNPMYCLPEYAGKDY 579
438 MLQINPDSSINPDHLSYFHFVGRIMGLAVPHGYINGGTVFPYKOLGKPIQLSDLESV 497
580 CLQINPASYINPDHLKYFRIGFIAMALPHGKFDITGSLFYKRLINKPVLKDLSEI 639
498 DPELHKLVLWILENDITPV--LDHTFCVHNAPGRILQHELKPNRNPVTEENKKEYVRL 556
640 DPEYNSLIWYXENIBECDEMYFSYDKELGEIKSHDLKPNNGNIIIVTEENKEYIRM 699
557 YVNRWVRGIEAOFALQKGFNELIPHLLKPPDQKELELIIGLDKLDNDKNTRLK 616
700 VAEWRLSGVEEQTAFFEGFNEILQQYLQYFDAKELEVLGCGQIEDLNDWQHAIR 759
617 HCVADSNVFWQAVETFDERRARLLQFVTGSTRVPLQGFKALQGSTGAAGPRLFTIH 676
760 HYARTSQIMKWFQVKEIDNEKRMELLOFVTGTCRLPVGGFADLNGSN---GPQKFCIE 816
677 LIDANTDNLKANTCNRIIDIPYESYKLYEKLTAVEETCGFAVE 723
817 KV-GKENWLPRTSHTCNRLDLPYKSYEQLEKLLPAIBETEGFQGE 862

;SULT 9
;GFE6
;Q96F66 PRELIMINARY; PRT; 303 AA.
;Q96F66
;C1-DEC-2001 (TrEMBLrel. 19, Created)
;C1-DEC-2001 (TrEMBLrel. 19, Last sequence update)
;01-MAR-2003 (TrEMBLrel. 23, last annotation update)
;Similar to Itchy (Mouse homolog) E3 ubiquitin protein ligase.
;Homo sapiens (Human).
;Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
;Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
;NCBI_TaxID=9606;
;[1]
;SEQUENCE FROM N.A.
;Tissue=Kidney;
;Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
;-1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
;EWBL; BC01571; AAH1571.1; -.
;Genew; HGNC:13890; ITCH.
;InterPro; IPR000008; C2.
;InterPro; IPR000569; HECT_domain.
;InterPro; IPR002349; WW_DOMAIN_2;
;InterPro; IPR001202; WW_Rap5_WWP.
```

```
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS00020; WW_DOMAIN_2; 2.
SQ SEQUENCE 903 AA; 102802 MW; 6777A2043C7B67BC CRC64;

Query Match 33.2%; Score 1288.5; DB 4; Length 903;
Best Local Similarity 33.8%; Pred. No. 5.5e-97;
Matches 312; Conservative 128; Mismatches 23; Indels 245; Gaps 29;

QY 6 KIRLTLYLCAK-NLAKKDFRPLDPFAKIVDVGSGCHSTDTVKNTLDPKMNQHYDLVYVK 64
Db 19 QLQITVISAKLKENKQWFG-PSPYVEVTVD--GQSKTEKKNNTNSPKWKQPLTVIVTP 75
QY 65 TSITISIVNHHKIHK-----KQAGFLG--88
Db 76 VSKLHPRVMSHOTLKSVDLLGTAALDIYETLKSNNMKLEEVVVTQLGDKKEPTTIGDL 135
QY 89 --CV-----RLSNALSRKADTGYQ--RDL-----CKL-----NPSDADAV 121
Db 136 SICDLGLQLESEVWINGETTCSENGVSLCLPRLECNLSAISAHCLCLPGLSDSPISARV 195
QY 122 RQIIVS-----LQTRD--RIQTGS-----VVDKRL-----LENGTVYEDSGP 160
Db 196 AGFTGASQNDGSRKDETRVSTNGSDPEDAGAGENRRVSGNNSPSLSNGG--FKPSRP 253
QY 161 GRPLSCFMEEPAPYTDSTGAAGGNCRFVES-----192
Db 254 PRP-----SRPPPTPRPASVNGSPSATSSESDSGSLDPTNTINTSGATSGLIIP 308
QY 193 -----PSQDRLQAQLRNPVDVRSGL-----QTPQNRPHGH 223
Db 309 LITSGSGRPLNPVTQAPLPPGWEQV-----DQGRVYVVDHVEKTTWDRPE--358
QY 224 QSPELPEGVEQRTTQGVYFLHTQTGVSTWHDPRIP-----260
Db 359 ----PLPPGWERVDNMGRIIYVDHFTRTTWTQRPFLSRYNVEQWLQORSOLQAGMOQFN 415
QY 261 -----RDL-----NSVNCDELGPLPPGWEVSTVSGRIYFVDHNNRTTQDRLHHIM 309
Db 416 QRFYGNQDLFATSSQKEFDPLGPLPGWEKRTDSNGRVIYFNENTRITQWEDR-----470
QY 310 NHQCLKEPSQPLPL-----PSGSELEDELPQAQYERD 343
Db 471 -SQGQLE--KPLPEGWEMRFTVDGIFYFVDHNNRTTYPDRTGKSALDNGPOIAYVRD 527
QY 344 LVQKLKVLRL---HELSQLQPOAGHCHRIEVSREEIFESYROIWKMRPKDLKRLMVKFRG 400
Db 528 FKAQVQYFRFWCQOLAMPQ---HIKITVTRKTLFESFQIINSFSDRLRRLLWVFP 583
QY 401 EBGLDYGGVAREWLYLLCHEMLNPPYGLFQYSTDNIIYMLQINPDSSINPDHLSYFHPVGR 460
Db 584 EBGLDYGGVAREWFLLSHEVLNPMYCLPEYAGKDYCLQINPASYINPDHLKYFRFTRIGR 643
QY 461 IMGLAVFHGYINGGTVFPYKOLGKPIQLSDLESVDPELHKLVLWILENDITPV-LDH 519
Db 644 FAMALFHGKFDITGFSIPFYKRLINKPVGLKDLSEIDPEFYNSLIWYKNNTEECDELM 703
QY 520 TFCVHNAPGRILQHELKPNRNPVTEENKKEYVRLYVNRWVRGIEAOFALQKGFNE 579
Db 704 YFSVDKEILGEIKSHDLKPNNGNIIIVTEENKEYIRMVAEWRLSRGVEEQTAFFEGFNE 763
QY 580 LIPOHLLKPFQDQKELELIIGLDKLDNDKNSNTRLKHCVADSNIVFWQAVETPDEER 639
Db 764 ILPQYLVQFDAKELEVLGCGQIEDLNDWQHAIRYHARTSQIMKWFQVKEIDNEK 823
```


Ingham R., Ernberg I., Pawson T.;
"Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases".
Mol. Cell. Biol. 20:8526-8535 (2000).
-!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
EMBL: AB071179; BAB69424.1; -.
InterPro: IPR000008; C2;
InterPro: IPR000569; HECT_domain.
InterPro: IPR002349; WW.
InterPro: IPR001202; WW_Reps_WWP.
Pfam: PF00168; C2; 1.
Pfam: PF00632; HECT; 1.
PRINTS: PR00360; C2DOMAIN.
PRINTS: PR00403; WWDOMAIN.
SMART: SM00239; C2; 1.
SMART: SM00119; HECT; 1.
SMART: SM00456; WW; 4.
PROSITE: PS00499; C2_DOMAIN_1; 1.
PROSITE: PS00004; C2_DOMAIN_2; 1.
PROSITE: PS00237; HECT; 1.
PROSITE: PS01159; WW_DOMAIN_1; 4.
PROSITE: PS00020; WW_DOMAIN_2; 4.
Ligase.
SEQUENCE 955 AA; 110021 MW; A8BB278A37F6A6B5 CRC64;

Query Match
Best Local Similarity 32.1%; Score 1247; DB 4; Length 955;
Matches 310; Conservative 155; Mismatches 246; Indels 292; Gaps 24;

2 GSSIKRLTVLCAKNAKDFRLPDPFAKI---VVDGSGQCH--STDVTKNTLDPKWNQ 56
17 GESRIILRVKVSIGDLAKKDFGASDYYVKLSLYADENRELALVQTKTKTLNPKWNE 76
57 HYDLVYKTD-STISVWNHKKHKQAGFLGCVRLLSNAI----- 97
77 EFTFRVNSNHRLLFEVDENRLTRD---FLGQVDVPLSLHFTEDPTMERYPTFKDFLL 133
98 -----SRKNDTGQYRLDCKLNPSDVAVRQIVVLSQTRDRTGSGSVVDCRG----- 146
134 RPRSHKSRVK-GFLKXWYMPKNG-----QDEENSQDQDDMEHGEVVDNSASQH 186
147 -----LLENEGTVY----- 155
187 QEELPPPLPPGWEKVDNLGRYYVNNHNRITQWRFSLMDVSSDNNRQINQEAH 246
156 -----EDSGP-----GRPLSCFMEEPAPYTDSTGAA-----AGGNCRFVES 192
247 RPRSRHHSEDLPPSPSGGVDPEFWETISEVNIAGDGLALPPPPASPGRSTPQE 306
193 PQSD-----QRLQARLN----- 206
307 LSEELSRRLQITPDSNGEQFSSLIQREPSRLRSCSVTDVAEQLHPPPSVAVVHTTPG 366
207 -----PDVR----- 210
367 LPSGWEERKAKGRYYVNNHNRITTTWTPINQLAEDGASGATNSNNHLEIPQIRPRS 426
211 -----GSLQTPQNRPHGHOSPE-----LPEGEYQR 235
427 LSSPTVTLAPLEGAKDSPVRAVVDTLNPSQSPSPNSPKQHKYQSLFPPGEMR 486
236 TTVQGVYFLHTQVSTWHDPRIP-----RLNSVNCDELGLPPGWEVSTVSGRIYF 290
487 IAPNGRPFFIDHNKTKTTWEDRLPKFPVHMRSKTSLNPNGLPLPPGWEERHLDGRTFY 546
291 VDNHNRITQTPDRLHHIMHQCQLEKPSQPLPLPSEGSLEDEELPAQRYERD.VQKLKV 350
547 IDHNSKITWEDRLQN-----PAITG-----PAPVYSEKQKVDY 583
351 LRHLSLQQAQAGCHRIEVSREIIEESYRQIMKM-RPKDLKKRLMWKFRGEGLDYGV 409
584 FRKLLKKPADIPNRFMKLHRNNIPESYRRYRWSVKRPDLVKARLWIEFESEKGLDYGV 643

```
157 PE-----PLPPGKERRVDNMGRIYYVDHFTTTTWTQRTLESVRYNQWLQORSQLOAM 211
261 -----RDL-----NSVNCDELGLPPLPGWEVRSTVSGRIYFVDHNNRTTQFTDRL 305
212 QQFNQRFYIGNQDLFATSQKSFDFPLPPLPGWEKRTDSNGRVYFVNHNTITOWEDPR- 270
306 HHIMNHQCCOLKEPSQPLPL-----PSEGSLEDEELPAQR 339
271 -----SQQQLNE--KPLPGWEVRFTVDGIPYFVDHNNRTTIDPRTGKSALDNGFQIA 323
340 YERDLVQKLKVLK---HELSLQOQAGHCRIEVSREBEIFESYRQIMKMRPKDLKRLMW 396
324 YVRDFKAKQYFRFWCQQLAMPQ---HIKITVTRKTLFEDSFQOIMSFSQDLRRRLMW 379
397 KPRGEGLDYGGVAREWLYLLCHEMLNYPVGLFOYSTDNIMYLOINPDSINPDHLSYFH 456
380 IFPEEGLDYGGVAREWFFLLSHEVLNPMYCLFAYAGKNTVCLQINPASYINPDHLKIFR 439
457 FVGRIMGLAVFHGHYINGGFTVPFYKQLGKPIQ-SDLESVDPELHKLVLWILENDITPV 516
440 FIGRFIAMALFHGKFIDTGSPLPFYKRLNKPVGLKLESIDPEFYNSLIWKENNIEEC 499
517 -LDHTFCVEHNAFRILQHELKXNGENVPVTEENKKEYVRLYVNWRFMGIEAQFLALQK 575
500 DLEMVFSVDKEILGEIKSHDLKPNNGNLTVEENKKEYIRMAEWRLSRGVEBOTQAFPE 559
576 GFNELIPQHLKPFQDKLELIGGLDKIDLNDWKSNTRLKHCVADSNIVRWFQAVET 635
560 GFNEILPQVLYFDAKELEVLVLCGMEIDLNDWQRIYRHARTSKQIWMFQVKEI 619
636 DEERRARLLQFVTGSTRVLPQFKALQSGTGAAGPRLFTIHLIDANTDNLPKARTCFNRI 695
620 DNEKRVLLQFVTGTCRLPVGGFADLMGNSN---GPKQFCIEKV-GKENWLPRSHTCFNRL 675
696 DIPPYSEYKLEKILTAVEETCGFAVE 723
676 DLPPYKSEYQLEKLLFAIEETEGFGQE 703

O43584 PRELIMINARY; PRT; 739 AA.
3584
SEQUENCE FROM N.A.
MEDLINE=98313405; PubMed=9647693;
Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
Kaminsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.:
"Acrophin-1, the DRPLA gene product, interacts with two families of WW
domain-containing proteins.";
Mol. Cell. Neurosci. 11:149-160 (1998).
EMBL: AF038564; AAC04845.1; -
HSP; Q13526; IPIN.
InterPro: IPR000569; HECT_domain.
InterPro: IPR002349; WW.
InterPro: IPR001202; WW_Rsp5_WWP.
Pfam: PF00632; HECT; 1.
Pfam: PF00397; WW; 4.
PRINTS: PR00403; WWDOMAIN.
SMART: SM00119; HECTC; 1.
SMART: SM00456; WW; 3.
PROSITE: PS50237; HECT; 1.
PROSITE: PS01159; WW_DOMAIN; 1; 4.
PROSITE: PS50020; WW_DOMAIN_2; 2.
```

```
FT NON_TER 1 1
SQ SEQUENCE 739 AA; 85057 MW; 7F1B4BASC105951B CRC64;
Query Match 32.1%; Score 1245.5; DB 4; Length 739;
Best Local Similarity 39.8%; Pred. No. 1.4e-93;
Matches 274; Conservative 99; Mismatches 180; Indels 135; Gaps 18;
QY 114 NPSDTDAVRGQIVVSL-----QTRDRIG-TGGSVVDCRGLLENEGTVYEDSGRPLSCF 167
DB 109 SFSATSESDGSGTSLSPPTNTNTSEGATSLIPL-----TISGSGP-RPLNPV 159
QY 168 MEEPAPYTDSTGAAAGGNCRFVESPQDQRLQARLNPDPVRGSL-----QTPOQR 219
DB 160 TQAPLP-----PCWEQRV-----DOHGRVYVDHVEKRTTWR 192
QY 220 PHGQSPPLPGYQRRTTVQGVYFHTQTGVSTHDPRI----- 260
DB 193 PE-----PLPPGKERRVDNMGRIYYVDHFTTTTWTQRTLESVRYNQWLQORSQLOAM 247
QY 261 -----RDL-----NSVNCDELGLPPLPGWEVRSTVSGRIYFVDHNNRTTQFTDRL 305
DB 248 QQFNQRFYIGNQDLFATSQKSFDFPLPPLPGWEKRTDSNGRVYFVNHNTITOWEDPR- 306
QY 306 HHIMNHQCCOLKEPSQPLPL-----PSEGSLEDEELPAQR 339
DB 307 -----SQQQLNE--KPLPGWEVRFTVDGIPYFVDHNNRTTIDPRTGKSALDNGFQIA 359
QY 340 YERDLVQKLKVLK---HELSLQOQAGHCRIEVSREBEIFESYRQIMKMRPKDLKRLMW 396
DB 360 YVRDFKAKQYFRFWCQQLAMPQ---HIKITVTRKTLFEDSFQOIMSFSQDLRRRLMW 415
QY 397 KPRGEGLDYGGVAREWLYLLCHEMLNYPVGLFOYSTDNIMYLOINPDSINPDHLSYFH 456
DB 416 IFPEEGLDYGGVAREWFFLLSHEVLNPMYCLFAYAGKNTVCLQINPASYINPDHLKIFR 475
QY 457 FVGRIMGLAVFHGHYINGGFTVPFYKQLGKPIQ-SDLESVDPELHKLVLWILENDITPV 516
DB 476 FIGRFIAMALFHGKFIDTGSPLPFYKRLNKPVGLKLESIDPEFYNSLIWKENNIEEC 535
QY 517 -LDHTFCVEHNAFRILQHELKXNGENVPVTEENKKEYVRLYVNWRFMGIEAQFLALQK 575
DB 536 DLEMVFSVDKEILGEIKSHDLKPNNGNLTVEENKKEYIRMAEWRLSRGVEBOTQAFPE 595
QY 576 GFNELIPQHLKPFQDKLELIGGLDKIDLNDWKSNTRLKHCVADSNIVRWFQAVET 635
DB 596 GFNEILPQVLYFDAKELEVLVLCGMEIDLNDWQRIYRHARTSKQIWMFQVKEI 655
QY 636 DEERRARLLQFVTGSTRVLPQFKALQSGTGAAGPRLFTIHLIDANTDNLPKARTCFNRI 695
DB 656 DNEKRVLLQFVTGTCRLPVGGFADLMGNSN---GPKQFCIEKV-GKENWLPRSHTCFNRL 711
QY 696 DIPPYSEYKLEKILTAVEETCGFAVE 723
DB 712 DLPPYKSEYQLEKLLFAIEETEGFGQE 739

RESULT 14
Q9NT88 PRELIMINARY; PRT; 820 AA.
AC Q9NT88;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (fragment).
GN DKF2P434P2422
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RS TISSUE=Testis;
RC Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
```

Submitted (JAN-2000) to the EMBL/GenBank/DBSJ databases.

-1- SIMILARITY: CONTAINS 1 C2 DOMAIN.

EMBL; AL137469; CAB70754.1; -

HSP; Q13526; 1PIN.

InterPro; IPR000008; C2.

InterPro; IPR000569; HECT_domain.

InterPro; IPR002349; WW.

InterPro; IPR001202; WW_Rsp5_WWP.

Pfam; PF00168; C2; 1.

Pfam; PF00632; HECT; 1.

Pfam; PF00397; WW; 3.

PRINTS; PR00403; WWDOMAIN.

SMART; SM00239; C2; 1.

SMART; SM00119; HECTC; 1.

SMART; SM00456; WW; 3.

PROSITE; PS00004; C2_DOMAIN_2; 1.

PROSITE; PS0237; HECT; 1.

PROSITE; PS01159; WW_DOMAIN_1; 3.

PROSITE; PS00020; WW_DOMAIN_2; 3.

Hypothetical protein.

NON TER

SEQUENCE 820 AA; 95283 MW; 0FDB34B29B3F4123 CRC64;

Query Match 32.0%; Score 1241.5; DB 4; Length 820;

Best Local Similarity 34.9%; Pred.No. 3.5e-93;

Matches 294; Conservative 116; Mismatches 230; Indels 203; Gaps 21;

43 TDVTKNTLPKQNHVDLVGKTD-SITISVNNKHKIKKQAGFLGCVRLLSNAI----- 97

12 TKTIKTLAPKNEEYFVFNSENHLLFEVDENLRTRDD---FLGQVDVPLSHLTED 68

98 -----SRKGTGYQRDLCKLNPSDSTDVAVGQIVVSLQTRDRIGT 137

69 PTMERPYTFKDFLLRPRSHKSRVK--GFLRLKWAYMPKNG-----GQDEENSQDQDDMEH 121

138 GGSVDPDRCG-----LLENEGTVY----- 155

122 GWEVDSNDASQHQELPPPLPPQWSEKVDNLGRTYYVNNHNTTQWHRPSLMDVSSE 181

156 -----EDSGP-----GRPLSCFMEEPAPYTDSTGAA-- 181

182 SDNNIRIQNEAAHRRFRGRRHISELEPESEGGVDPPEMETISEEVNIAGDSLGLALP 241

182 ---AGGNCRFVESPDQ-----ORLQAQRLN----- 206

242 PFPASPGRSTSPQELSEELSRRLQITPDNGSQFSLQIRESSRLRSCSVTDAVEQGH 301

207 -----PDVRGSLQTPQNRPHGHQSPE-----LPEGYQRITVQGVYFL 245

302 LPPGAKDSPVRAVKDTLSNPQSPSPYNSPKQKHVQTSFLPPGWMRIAPNGRPFI 361

246 HTQTGVSTWHDPRIP-----RDLSNVNCDLGLPLPGWEVRSVTSGRIFYVDHNNRTQF 300

362 DHNTKTWTDPRLKPFPVHRMRSKTSLNPNLGLPLPGWEERIHLDGRTFYFDHNSKITQW 421

301 TDPLRLHMHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLVQKLVRHESLSQOP 360

422 EDPLQN-----PAITG-----PVPYSEFPKQDYFKKLKKPAD 458

361 QAGHCRLEVRSEEIFEESYRQIMM-RPKDLKRLMKVFRGEGLDYGVAREWLYLCH 419

459 IPNRFENKLRHNNIFEESYRIRINSVKRPDVLKARLWIEPESEKGLDYGAVAREWFLLSK 518

420 EMLNPFYGLQYSDNIMYLQINPDSSI-NPDHLSYFVFVGRIMGLAVPHGYINGGFTV 478

519 EMFNPFYGLFEYATNITLQINPNGLCNEDHLSYFTIGRVAGLAVPHGKLLDGFIR 578

479 PFYKQLGKPIQLSDLESVDPELHSLWILENDITPVLDTFCVNEHNAFGRILQHELP 538

579 PFYKWMGLGKQITLNDMSVDSEYNSLKWILENDPTE-LDLAFCDIEENFGQTYQVDLKP 637

539 NGRNVPTTEENKEVYVLYNTEFMRGIEAQFLALQKGFNELIPQHLKPPDQKELELI 598

Db 638 NGSELTWNTENKREYIDLVIQWRVNRVQKMAFLEGTELLPIDLIKIFDNELELLM 697

QY 599 GLDKIDLDNWKSNRLKHCVADSN-IVRWFQAVETFEDEERARLLQFVTGSTRVPLOG 657

Db 698 CGLGDVDNDWRQHSIYKNGYCPNHPVQWFWKAVLLMDAEKRIQLQFVTGSTRVPMNG 757

QY 658 FKALGGSTGAAGPRLFTHLIDANTDLNPKAHTCFNRIDIPPYSEYKLYEKLTTAVEET 717

Db 758 FAELYGSN---GPQLFTIEQW-GSPEKLPRAHYCFNRLDLPFYETFEDELEKLLMAVENA 813

QY 718 QGF 720

Db 814 QGF 816

RESULT 15

Q9HOMO PRELIMINARY; PRT; 922 AA.

AC Q9HOMO; 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Hypothetical protein (WW domain-containing protein 1).

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=21154917; PubMed=11230166;

RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,

RA Ansoerg W., Boecher M., Bloecher H., Bauersachs S., Blum H.,

RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,

RA Mewes H.W., Othenwaelder B., Obermaier B., Tampe J., Heubner D.,

RA Wambutt R., Korn B., Klein M., Poustka A.;

RT Towards a Catalog of Human Genes and Proteins: Sequencing and

RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.;

RL Genome Res. 11:422-435(2001).

RN [2]

RP SEQUENCE FROM N.A.

RA Flasz M., Aslam H., Poulson R., Hanby A., Gorman P., Roylance R.,

RA Nolan T., Brady G., Canfield A., Baron M.;

RT "WPI", a human homolog of Drosophila suppressor of deltex.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA Strausberg R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.

CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.

DR EMBL; AL137469; CAB66673.1; -

DR EMBL; AY043361; RAK9468.1; -

DR EMBL; BC036065; RAK36065.1; -

DR HSP; Q13526; 1PIN.

DR InterPro; IPR000008; C2.

DR InterPro; IPR000569; HECT_domain.

DR InterPro; IPR002349; WW.

DR InterPro; IPR001202; WW_Rsp5_WWP.

DR Pfam; PF00168; C2; 1.

DR Pfam; PF00632; HECT; 1.

DR Pfam; PF00397; WW; 4.

DR PRINTS; PR00403; WWDOMAIN.

DR SMART; SM00239; C2; 1.

DR SMART; SM00119; HECTC; 1.

DR SMART; SM00456; WW; 4.

DR PROSITE; PS00004; C2_DOMAIN_2; 1.

DR PROSITE; PS0237; HECT; 1.

DR PROSITE; PS01159; WW_DOMAIN_1; 4.

DR PROSITE; PS00020; WW_DOMAIN_2; 2.

DR Hypothetical protein.

KW SEQUENCE 922 AA; 105201 MW; 35B6E1C03A3147DA CRC64;

```

Query Match      31.9%, Score 1239, DB 4; Length 922;
Best Local Similarity 32.5%, Pred.No.6.8e-93;
Matches 300; Conservative 138; Mismatches 262; Indels 224; Gaps 25;

5 IKIRLTVLCANLAKKOPFRLPDPFAKIVVDGSGQCHSTDTVK--NTLDPKMNQHYDLV 62
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18 LQLQVTVSSAKLKKKWFQTA-IYTEVVVG---EITKAKSSSSNPKNDEQLTVNV 72
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 GKTDSITISVNHKKIHKQAGFLG-----CVRLLSNALSLKD----- 102
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
73 TPQTTLEFQVWSHRTL---KADALLGKATIDLKQALLIHNKLERVKEQKLSLENKNGI 129
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
103 --TG-----YORLDCXKNPS-----DTDAVRGQIVVSLQTRDRI---GTGG- 139
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
130 ACTGELTVVLDGLVIEQENINCSSSTIBIQENGDAHENGEPSSARTTARLAVEGTNGI 189
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
140 -SVDRCGLLENEGTVYEDSGFGRPLSCFM-----EEPAP-----Y 174
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
190 DNHVPTSLVQNSCCSYVVGNDTFSSPSQVAARPKNTPAPKPLASEPADDTVNGESSSF 249
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
175 TDSCTGAAAGG-----GNC-----RFVESPSODORLQAOR 203
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
250 APTDNASTVGTVPVSEENALSPNCTITVEDPPVQEIITSENNECIPSTSAEISEARS 309
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
204 LKNPDVRSLOT-----PONRPHGHOSP-----ELPEGVEQRTTVQGVYFLH 246
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
310 ILEPDTNSRSSSAFAEAAKSRQPCDCMPVRQSGNANTETLPSGWEQRKDPHGRYYVD 369
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 TOTGVSTWHDPR-IP----- 260
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
370 HNTRITTWERPQLPFGWERRVDRRRVYVDHNTRTTTWQRTMESVRNFEQWQSRNQ 429
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
261 -----RDLSVNC-----DELGLPFGWEVRSTVSGRIYFVDHNNRTTQTTDRL 305
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
430 LQAGMQCFNRQVLYSASMLAENDPYGLPGWEKRVDSIDRVYFVNHTKTTQWEDPRT 489
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
306 HIMMHO-----CQLKPSQPLPLPSEGSLEDELPQARY 340
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
490 QGLQNEEPLPEGWELRYTREGVRYFVDHNTTTTFKDPNGKSVTKG-----PQIAY 543
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
341 ERDLVKLKLRLSHLSLQPOAGHCHRIEVSREEIPESYRQIMQRPKDLKRLMVKFRG 400
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
544 ERGFRWKLAFHY-LCQSNALPSHVKNVSRQTLFEDSFQIIMALKPYDLRRRLYVIFRG 602
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
401 EBGLDYGVAREWLYLLCHEMLNPYYGLFOYSTDNIMQLINPDSSINPDHLSYFHFVGR 460
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
603 EBGLDYGLAREWFFLLSHEVLNPMYCLFEYAGKNYCLQINPASTINPDHLSYFCFGR 662
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
461 IMGLAVFHGHYINGGFTVPFYKOLLGKPIQLSDLESVDPELHKLVLWILENDITPV-LDH 519
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
663 FIAMALFHGKFIDTGFSLPFYKMLSKLTIKLESIDTFYNSLIWIRDNNEECGLEM 722
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
520 TFCVHNHAFGRILQHELKPNGRNVPVTENKKEYVRLVNNRFRMGIEAQFLAQGFNE 579
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
723 YFSDMEILGKVTSHDLKLGSSNIIIVTEENKDEYICGLMTEWRFSRGVQEQTKAFLDGFNE 782
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
580 LIPOHLKPFDDOKELIIGGLDKIDNDKNSNTRLKHCVADSNIVRWFQAVETDEER 639
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
783 VPLQWLQYFDEKELEVMCGQVEVDLADWQNTVYRHYTNSKQIIWFWQFVKETDNEV 842
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
640 RARLQFVTGSTRVPLQGFKALQSGTGAAGPRLETHLIDANTDNLPKAHTCFNRIDIPP 699
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
843 RMELLQFVTGTCRLPLGGFAELMGSN---GPKFCIEKVGKDT-WLPSHTCFNRDLDP 898
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
700 YESYEKLYEKLITAVEETCGFAVE 723
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
899 YKSYEQLEKELFAIEBTEGFGQS 922
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

on: February 20, 2004, 15:25:26 ; Search time 41.6968 Seconds
(without alignments)
2847.400 Million cell updates/sec

le: US-10-009-945-4

ect score: 4038

ence: 1 MGNPGRRRGPKVRLTVLC.....EKLYEKLTAETTCGPAVE 748

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 1107863

imum DB seq length: 0

imum DB seq length: 200000000

b-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : A_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	Query Match	Length	ID	Description
1	4038	100.0	748	22	AA1980.DAT.*
2	4031	99.8	748	18	AA1981.DAT.*
3	4027	99.7	804	22	AA1982.DAT.*
4	3001	74.3	722	24	AA1983.DAT.*
5	3001	74.3	722	22	AA1984.DAT.*
6	2992	74.1	722	22	AA1985.DAT.*
7	2775	68.7	514	22	AA1986.DAT.*
8	2173	53.8	1035	22	AA1987.DAT.*
9	1545	38.3	766	18	AA1988.DAT.*

10	1530.5	37.9	832	23	ABP73459	Candida albicans e
11	1453.5	36.0	869	24	ABJ26104	Aspergillus fumiga
12	1387	34.3	975	20	AAW3167	Human ZGBPI prote
13	1355.5	33.6	724	18	AAW36797	Novel human gene,
14	1354.5	33.5	927	20	AAZ51170	Human KIAA ligase
15	1354.5	33.5	927	24	AAE32725	Human NEDD4 long f
16	1353.5	33.5	995	24	AAE32719	KIAA0439 protein.
17	1353	33.5	923	22	ABG01080	Novel human diagno
18	1348	33.4	834	23	AAU77715	Diseased kidney ei
19	1340.5	33.2	759	24	ABJ25504	Aspergillus fumiga
20	1338.5	33.1	854	24	ABP58332	Human cell growth,
21	1336	33.1	854	20	AAZ30949	Murine E3 ubiquiti
22	1333	33.0	949	22	ABP59631	Drosophila melanog
23	1333	33.0	949	22	ABP59856	Drosophila melanog
24	1329	32.9	852	20	AAZ30948	Human E3 ubiquitin
25	1301	32.2	739	22	ABG16477	Novel human diagno
26	1301	32.2	739	22	ABG68173	Atrophin-1 interac
27	1301	32.2	739	24	ABR41097	Human atrophin-1 i
28	1301	32.2	739	24	AAE32720	Atrophin-1 interac
29	1298.5	32.2	870	24	AAE05495	Human ubiquitin pr
30	1298.5	32.2	870	24	AAE32722	Nedd-4-like ubiqui
31	1297	32.1	922	22	AAE50049	Human homolog of D
32	1297	32.1	922	23	ABP5708	Human signal trans
33	1285.5	31.8	906	18	AAW36795	Novel human protei
34	1281	31.7	733	22	AAE50048	Human clonore 811a
35	1273	31.5	1071	22	ABG11772	Novel human diagno
36	1271	31.5	818	18	AAW13386	Human protein ubiq
37	1265.5	31.3	1082	22	ABG63355	Drosophila melanog
38	1209	29.9	474	20	AAZ10943	Amino acid sequenc
39	1134.5	28.1	683	18	AAW36794	Novel human protei
40	1134.5	28.1	683	22	AAE05494	Human ubiquitin pr
41	1134.5	28.1	684	24	AAE32721	Nedd-4-like ubiqui
42	1032	27.0	1572	23	ABP7562	Novel human protei
43	1072	26.5	335	22	AAW78877	Human protein SEQ
44	1050	26.0	1562	22	ABG7493	Novel human diagno
45	1050	26.0	1562	24	AAE32723	KIAA0322 protein.

ALIGNMENTS

RESULT 1

AA1980.DAT.*

ID AAB31477 standard; Protein; 748 AA.

AC AAB31477;

DT 20-APR-2001 (first entry)

XX Amino acid sequence of a human Smurf2 polypeptide.

DE Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP;
 KW transforming growth factor-beta; human; TGF-beta; chondrogenesis;
 KW osteogenesis; blood differentiation; cartilage formation; hair growth;
 KW neural tube patterning; retinal development; heart induction;
 KW morphogenesis; tooth formation; gamete formation.
 OS Homo sapiens.

XX WO20007168-A2.

PN 21-DEC-2000.

PD 12-JUN-2000; 2000WO-US16250.

PF 11-JUN-1999; 99US-0138969.

PR (UNY) UNIV NEW YORK STATE RES FOUND.

PA (HSCR-) HSC RES & DEV LP.

XX Thomsen GH, Wrana J;

XX WPI; 2001-071267/08.

Query Match	99.7%;	Score 4027;	DB 22;	Length 804;
Best Local Similarity	99.7%;	Pred. No. 0;	Mismatches 2;	Indels 0;
Matches 746;	Conservative 0;			
1	MSNPGRRRGPKVRLTLVLCANLVKQDFRLLPDPFAKVVVDGSGQCHSTDTVNTLDPK	60		
57	MSNPGRRRGPKVRLTLVLCANLVKQDFRLLPDPFAKVVVDGSGQCHSTDTVNTLDPK	116		
61	WQHGYDLYTGKSDSVTISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL	120		
117	WQHGYDLYTGKSDSVTISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL	176		
121	GPNDNDTVRGQIVVSLQSRDRIGTGQGVVDCSRLFDNDLPDGCWEERRTASGRIOYLNHIT	180		
177	GPNDNDTVRGQIVVSLQSRDRIGTGQGVVDCSRLFDNDLPDGCWEERRTASGRIOYLNHIT	236		
181	RTTQWERPRTPASEYSSGRPLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN	240		
237	RTTQWERPRTPASEYSSGRPLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN	296		
241	YMSRTHLTPPDLPEGYEORTTQOGQVYFLHTQTGVSTWHDPRVPRDLNSINCELGLP	300		
297	YMSRTHLTPPDLPEGYEORTTQOGQVYFLHTQTGVSTWHDPRVPRDLNSINCELGLP	356		
301	PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVSLCPDDT	360		
357	PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVSLCPDDT	416		
361	ECLTVPRYKEDLVQKLIKILROELSQQOQAGHCRIEVSREEIFEESYQVMKRPKDLWK	420		
417	ECLTVPRYKEDLVQKLIKILROELSQQOQAGHCRIEVSREEIFEESYQVMKRPKDLWK	476		
421	RLMIKFRGEGLDYGVAREWLYLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL	480		
477	RLMIKFRGEGLDYGVAREWLYLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL	536		
481	SYPHFVGRIMGMAVFRHGYIDGGFTLPFYKQLLGSITLDDMELVDPDLNSLWILEND	540		
537	SYPHFVGRIMGMAVFRHGYIDGGFTLPFYKQLLGSITLDDMELVDPDLNSLWILEND	596		
541	ITGVLDHTFCVHNAYGEIIOHELKPKNGKSIPVNEENKEYVRLYVNRFLRGIEAFLA	600		
597	ITGVLDHTFCVHNAYGEIIOHELKPKNGKSIPVNEENKEYVRLYVNRFLRGIEAFLA	656		
601	LQGFNEVIPQHLTKTFDEKELELIICGLGKIDVNDWKVNRFLKCHTSPDSNIVKWPFAV	660		
657	LQGFNEVIPQHLTKTFDEKELELIICGLGKIDVNDWKVNRFLKCHTSPDSNIVKWPFAV	716		
661	EFFDEERRARLLQFVTGSSRVPLQGFKALQGAAGPRLFTIHOIDACTNNLPAKHTCFNRI	720		
717	EFFDEERRARLLQFVTGSSRVPLQGFKALQGAAGPRLFTIHOIDACTNNLPAKHTCFNRI	776		
721	DIPYSEYKLYEKLLTAIEETCGFAVE	748		
777	DIPYSEYKLYEKLLTAIEETCGFAVE	804		
Query Match	74.3%;	Score 3001;	DB 24;	Length 722;
Best Local Similarity	74.6%;	Pred. No. 3.8e-274;		
Matches 565;	Conservative 63;	Mismatches 71;	Indels 58;	Gaps 8;
QY	12	VKLRTVLCAKNLVKQDFRLLPDPFAKVVVDGSGQCHSTDTVNTLDPKWNQHYDLYGK	71	
Db	4	IKIRLTVLCAKNLVKQDFRLLPDPFAKVVVDGSGQCHSTDTVNTLDPKWNQHYDLYGK	63	
QY	72	SPSVTISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKLGNPDNDTVRGQ	131	
Db	64	TDSITISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKLGNPDNDTVRGQ	123	
QY	132	IVVSLQSRDRIGTGQGVVDCSRLFDNDLPDGCWEERRTASGRIOYLNHITTTOWERPTRP	191	
Db	124	IVVSLQSRDRIGTGQGVVDCSRLFDNDLPDGCWEERRTASGRIOYLNHITTTOWERPTRP	154	
QY	192	ASEYSSGRPLSCFVDENTPISGTNGATCG-----QSSDPRLAERRVRSQRHNRYM	242	
Db	155	---EDSGRPLSCFVEEPAPYTDSTGAAGGNCGRFVESFSQDQRLQAQRLRNPVGRSL	212	
QY	243	----SRTHLTPPOLPEGYEORTTQOGQVYFLHTQTGVSTWHDPRVPRDLNSINCELGP	296	
Db	213	QTPQNRPHGHOSPELPEGYEORTTQOGQVYFLHTQTGVSTWHDPRVPRDLNSINCELGP	272	
QY	299	LPPGHEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVW----	354	
Db	273	LPPGHEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVW----	328	
QY	355	LCPDDETECLTVPRYKEDLVQKLIKILROELSQQOQAGHCRIEVSREEIFEESYQVMKRM	414	
Db	329	L---EDELPAQRYERDLVQKLIKILROELSQQOQAGHCRIEVSREEIFEESYQVMKRM	385	
Query Match	74.3%;	Score 3001;	DB 24;	Length 722;
Best Local Similarity	74.6%;	Pred. No. 3.8e-274;		
Matches 565;	Conservative 63;	Mismatches 71;	Indels 58;	Gaps 8;
QY	12	VKLRTVLCAKNLVKQDFRLLPDPFAKVVVDGSGQCHSTDTVNTLDPKWNQHYDLYGK	71	
Db	4	IKIRLTVLCAKNLVKQDFRLLPDPFAKVVVDGSGQCHSTDTVNTLDPKWNQHYDLYGK	63	
QY	72	SPSVTISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKLGNPDNDTVRGQ	131	
Db	64	TDSITISVWNHKKTHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKLGNPDNDTVRGQ	123	
QY	132	IVVSLQSRDRIGTGQGVVDCSRLFDNDLPDGCWEERRTASGRIOYLNHITTTOWERPTRP	191	
Db	124	IVVSLQSRDRIGTGQGVVDCSRLFDNDLPDGCWEERRTASGRIOYLNHITTTOWERPTRP	154	
QY	192	ASEYSSGRPLSCFVDENTPISGTNGATCG-----QSSDPRLAERRVRSQRHNRYM	242	
Db	155	---EDSGRPLSCFVEEPAPYTDSTGAAGGNCGRFVESFSQDQRLQAQRLRNPVGRSL	212	
QY	243	----SRTHLTPPOLPEGYEORTTQOGQVYFLHTQTGVSTWHDPRVPRDLNSINCELGP	296	
Db	213	QTPQNRPHGHOSPELPEGYEORTTQOGQVYFLHTQTGVSTWHDPRVPRDLNSINCELGP	272	
QY	299	LPPGHEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVW----	354	
Db	273	LPPGHEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQKLDKQDQOQVW----	328	
QY	355	LCPDDETECLTVPRYKEDLVQKLIKILROELSQQOQAGHCRIEVSREEIFEESYQVMKRM	414	
Db	329	L---EDELPAQRYERDLVQKLIKILROELSQQOQAGHCRIEVSREEIFEESYQVMKRM	385	

E3 ubiquitin ligase SMURF1 protein.

Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis; poliomyelitis; HIV; measles; protein therapy; E3 ubiquitin ligase; enzyme.

Unidentified.

415 PKDLWELMIKPEEGEGLYGGVAREWYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 474
 386 PKDLKRLMWKFGEGEGLYGGVAREWYLLSHEMLNPPYGLFOYSTONIMYQLINPDSA 445
 475 VNPHELSYFHFVGRIMGMAVFFHGHYIDGGFTLPFYKQLIGKSIITLDDMELVDPOLHNSLV 534
 446 INPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQLIGKPIQLSDLESVDPELHNSLV 505
 535 WILENDITGVLDHTFCVEHNAVEIIOHELKPNKSIIPVNEENKKEYVRLVYNNRFLRGI 594
 506 WILENDITPVLDHTFCVEHNAFGRILQHELKPNRNPVTEENKKEYVRLVYNNRFRMGI 565
 595 EAQFLALQKGFNEVIFQHLLKTPDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVK 654
 566 EAQFLALQKGFNELIFQHLLKTPDQKELELIIGGLDKIDLNDKSNTRLKHCVADSNIVR 625
 655 WFKVAVEFDEERARLLQFVTGSSRVPLQGFALQ---GAAGPRLFTIHQIDACTNNLP 711
 626 WFKQAVETFEERARLLQFVTGSTRVPLQGFALQGSTGAAGPRLFTIHLDANTNLP 685
 712 KAHTCFNRIDIPPESEYKLYEKLTAIEETCGFAVE 748
 686 KAHTCFNRIDIPPESEYKLYEKLTAIEETCGFAVE 722

ULT 5
 31476
 AAB31476 standard; Protein; 723 AA.

AAB31476;

20-APR-2001 (first entry)

Amino acid sequence of a human Smurf1 polypeptide.

Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP; transforming growth factor-beta; human; TGF-beta; chondrogenesis; osteogenesis; blood differentiation; cartilage formation; hair growth; neural tube patterning; retinal development; heart induction; morphogenesis; tooth formation; gamete formation.

Homo sapiens.

WO2000077168-A2.

21-DEC-2000.

12-JUN-2000; 2000WO-US162250.

11-JUN-1999; 99US-0138969.

(UTNY) UNIV NEW YORK STATE RES FOUND.
 (HSCR-) HSC RES & DEV LP.

Thomsen GH, Wrana J;

WPI; 2001-071267/08.

N-P8DB; AAF24852.

Novel isolated Smurf protein useful for inhibiting bone morphogenic protein or tumor growth factor-beta activation pathway, for treating cancer and to block osteogenesis, hair growth, tooth formation -

Claim 6; Fig 10; 107pp; English.

The present sequence represents a human Smurf1 polypeptide. The specification also describes a Smurf2 polypeptide. Smurf polypeptides are negative regulators of Smad signal transduction, and antagonists of bone morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta) signalling pathway. Expression of Smurf1 in a cell is useful for inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf polypeptides are useful for blocking chondrogenesis, osteogenesis, blood differentiation, cartilage formation, neural tube patterning, retinal

CC development, heart induction and morphogenesis, hair growth, tooth formation, gamete formation and a wide variety of tissue and organ formation processes, and hinder the regeneration, growth, maintenance, etc., of bone and other tissues that are dependent on the BMP pathway. The polypeptide is useful for screening for various drugs and/or antibodies that can either enhance the BMP pathway, or inhibit it.
 XX
 SQ Sequence 723 AA;
 Query Match 74.3%; Score 3001; DB 22; Length 723;
 Best Local Similarity 74.6%; Pred. No. 3.9e-274;
 Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;
 QY 12 VKLRITVLCARLVKVKDFRDPDFAKVVDGSGQCHSTDTVKNTLDPKKNQCHVDLYIGK 71
 DB 5 IKIRITVLCARLVKVKDFRDPDFAKVVDGSGQCHSTDTVKNTLDPKKNQCHVDLYIGK 64
 QY 72 SSVTISVNNKKIHKQAGFLGCVRLLSNAIRLKDGTGYQRLDCKLGNFNDNTRVQG 131
 DB 65 TDSITISVNNKKIHKQAGFLGCVRLLSNAIRLKDGTGYQRLDCKLGNFNDNTRVQG 124
 QY 132 IVVSLQSDRIGTGQGVVDCSRLFDNDLPDQWEERTASGRIOYLNHITRTQWERTRP 191
 DB 125 IVVSLQSDRIGTGQGVVDCRGLLENE-----GIVY----- 155
 QY 192 ABEYSPPGRPLSCFVDENTPISGTNGATCG-----QSSDPRLAERRVRSQRHRYM 242
 DB 156 --EDSGPKRPLSCFMEEPAPVTDSTGAAAGGNCRFVESPQDQRLQAQLRNPVDRGSL 213
 QY 243 ----SRTHLTPDLPPEGYEORTTCOGVYELHTQTGVSTWHDPRVPRDLSNCELGP 298
 DB 214 QTFQNRPHGHQSPPELPEGYEORTTCOGVYELHTQTGVSTWHDPRVPRDLSNCELGP 273
 QY 299 LPPGWEIRNTATGRVYFVDHNNRTTQFDPRLSANLHLVLRNQQLKXQQQV---S 354
 DB 274 LPPGWEIRNTATGRVYFVDHNNRTTQFDPRLSANLHLVLRNQQLKXQQQV---S 329
 QY 355 LCPDDTECLTVPRYKRLVQKLKILRQLSQQQOAGHCRIEVSREIPEESYRQVMKR 414
 DB 330 L---EDEELPAQRYERDLVQKLKVLRLHELSQLQOAGHCRIEVSREIPEESYRQVMKR 386
 QY 415 PKDLWELMIKPEEGEGLYGGVAREWYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 474
 DB 387 PKDLKRLMWKFGEGEGLYGGVAREWYLLSHEMLNPPYGLFOYSTONIMYQLINPDSA 446
 QY 475 VNPHELSYFHFVGRIMGMAVFFHGHYIDGGFTLPFYKQLIGKSIITLDDMELVDPOLHNSLV 534
 DB 447 INPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQLIGKPIQLSDLESVDPELHNSLV 506
 QY 535 WILENDITGVLDHTFCVEHNAVEIIOHELKPNKSIIPVNEENKKEYVRLVYNNRFLRGI 594
 DB 507 WILENDITPVLDHTFCVEHNAFGRILQHELKPNRNPVTEENKKEYVRLVYNNRFRMGI 566
 QY 595 EAQFLALQKGFNEVIFQHLLKTPDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVK 654
 DB 567 EAQFLALQKGFNELIFQHLLKTPDQKELELIIGGLDKIDLNDKSNTRLKHCVADSNIVR 626
 QY 655 WFKVAVEFDEERARLLQFVTGSSRVPLQGFALQ---GAAGPRLFTIHQIDACTNNLP 711
 DB 627 WFKQAVETFEERARLLQFVTGSTRVPLQGFALQGSTGAAGPRLFTIHLDANTNLP 686
 QY 712 KAHTCFNRIDIPPESEYKLYEKLTAIEETCGFAVE 748
 DB 687 KAHTCFNRIDIPPESEYKLYEKLTAIEETCGFAVE 723

RESULT 6
 AAM79861
 ID AAM79861 standard; Protein; 722 AA.
 XX
 AC AAM79861;
 XX
 DT 06-NOV-2001 (first entry)

Human protein SEQ ID NO 3507.	Db	155	---EDSGEGRPLSCFMEEPAPYTDSTGAAGGNCRFVESPQDQRLQORLANPDVRSGL	212
Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation.	Qy	243	---SRTHLTPDLPEGEQRTTQGGVYFHTQTGVSTWHDPRVPRDLNINCEELGP	298
	Db	213	QTPQNRPHGHOSBELPEGEQRTTQGGVYFHTQTGVSTWHDPRVPRDLNINCEELGP	272
	Qy	299	LPGWEIRNTATGRVYFVDHNETTQFTDPRLSANLHLVLAERONOLKQDQDQVW---S	354
Homo sapiens.	Db	273	LPGWEVRSTVSGRIYFVDHNRRTTQFTDPR---	328
WC200157190-A2.	Qy	355	LCPDTECLTVPRYKRDLYQKILRQELSOQPGAGHCRIEVSREEIFEESYQVMKOR	414
09-AUG-2001.	Db	329	L---EDESLPAQYERDLVQKLVLRHLSLQPPQAGHCRIEVSREEIFEESYQVMKOR	385
05-FEB-2001; 2001WO-US04098.	Qy	415	PKLWKLMIKFRGEGGLDYGGVAREWLYLLSHEMLNPYGLFOYSRDDIYTLQINPDA	474
03-FEB-2000; 2000US-0496914.	Db	386	PKDLKGLMVKFRGEGGLDYGGVAREWLYLLCHEMLNPYGLFOYSTONIMQLINPDS	445
27-APR-2000; 2000US-0560875.	Qy	475	VNPEHLSYFHFVGRIMGMVFFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHSLV	534
20-JUN-2000; 2000US-0598075.	Db	446	INPDLSYFHFVGRIMGLAVFHHGHYINGGFTVPFYKQLLGKPIQLSDLESVDPELHSLV	505
19-JUL-2000; 2000US-0620325.	Qy	535	WILENDITGVLDHTFCVEHNAYGEIIQHELKPKNGKSI PVNEENKKEYVRLYNNWFLAGI	594
01-SEP-2000; 2000US-0654936.	Db	506	WILENDITPVLDHTFCVEHNAYGEIIQHELKPKNGRNPVTEENKKEYVRLYNNWFLAGI	565
13-SEP-2000; 2000US-0663561.	Qy	595	BAQFLAQKGFNEVPOHLLKTFDEKELELIICGLKIDVNDKVNTRLKHCTPDSNIVK	654
20-OCT-2000; 2000US-0693325.	Db	566	BAQFLAQKGFNELIPOHLLKTFDEKELELIIGGLKIDLNDKSNTRLKHCVADSNIVR	625
30-NOV-2000; 2000US-0728422.	Qy	655	WFKXAVEFFDEERRARLLQFVTGSSRVLPQGGFKALQ---	711
(HYSE-) HYSEQ INC	Db	626	WFMQAVETFEERRARLLQFVTGSTRVPLQGGFKALQSGTGAAGPRLFTIHLIDANTDLR	685
Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y; Zhao QA, Wang P, Wang J, Zhang J, Ren F, Chen R, Wang ZH; Xue AJ, Yang Y, Wejhrman T, Goodrich R;	Qy	712	KAHTCFNRIDIPPVESYEKLYEKLTAIETCGFAVE	748
WPI; 2001-476283/51.	Db	686	KAHTCFNRIDIPPVESYEKLYEKLTAIETCGFAVE	722
N-PSDB; AAK52994.	Qy			
Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy -	Db			
Claim 20; Page 362-363; 6221pp; English.				
The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activity/inhibit activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.				
Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.				
Sequence 722 AA,				
ery Match				
st Local Similarity				
atches 564; Conservative 63; Mismatches 72; Indels 58; Gaps 8;				
12 VKLRLTVLCAKLVKDKFRLPDPFPAKVVVDGSGGCHSTDTVNTLDPKMNQHYDLVIGK				71
4 IKRLVLCALAKLAKDFRLPDPFPAKVVVDGSGGCHSTDTVNTLDPKMNQHYDLVIGK				63
72 SDSVTISVNNKKIHKKQAGFLGCVRLLSNINRLKDTGTQRLDCLKGPNNDTVRGQ				131
64 TDSITISVNNKKIHKKQAGFLGCVRLLSNINRLKDTGTQRLDCLKGPNNDTVRGQ				123
132 IVVSLOSRRDRIGTGGGVDCSRFLFNDLDPDGWEERTASGRIOYLNHITRTQWERTRP				191
124 IVVSQQRDRIGTGGGVDCRGLLE-----GIVY-----				154
192 ASEYSFGFRPLSCFVDENTPISGTNGATCG-----QSSDPRLAERVRVSORHRYM				242

The invention describes an isolated nucleic acid molecule (1) encoding a novel central nervous system protein. (1) and polypeptides (11) encoded by (1), are used to treat a medical condition and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malformations, respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial infarction. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 68.7%; Score 2775; DB 22; Length 514;
 est Local Similarity 99.8%; Pred. No. 5.2e-253;
 atches 513; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

235 SQRHNYMSRTHLTPDLPEGEQRTQGGVYFLHTGTGVSTWHDPRVRLNINCE 294
 1 SRRHNYMSRTHLTPDLPEGEQRTQGGVYFLHTGTGVSTWHDPRVRLNINCE 60

295 ELGPLPGEWIRNTATGRVYFVDHNRRTFTDPRLSANLHLVLRNQLKDDQQQQVVS 354
 61 ELGPLPGEWIRNTATGRVYFVDHNRRTFTDPRLSANLHLVLRNQLKDDQQQQVVS 120

355 LCPDTECLTVPRYKRLVOKLTLRQLSQQPOQAGHCRIEVSREEFEESYRQVMKMR 414
 121 LCPDTECLTVPRYKRLVOKLTLRQLSQQPOQAGHCRIEVSREEFEESYRQVMKMR 180

415 PKDLWKRLMKFRGEGLDYGAVREWLVLSSHEMLNPNYGLFYQSRDDIYTLQINPDSA 474
 181 PKDLWKRLMKFRGEGLDYGAVREWLVLSSHEMLNPNYGLFYQSRDDIYTLQINPDSA 240

475 VNPBLSVFFHVRIMGMAVPHGYIDGGFTLPFYKQLGKSTLDDMLVDPDLNLSLV 534
 241 VNPBLSVFFHVRIMGMAVPHGYIDGGFTLPFYKQLGKSTLDDMLVDPDLNLSLV 300

535 WLENDITGVLDHTFCVEHNAYGRIIOHELKPNKSTPVNEENKKEYVRLYNWFLRGI 594
 301 WLENDITGVLDHTFCVEHNAYGRIIOHELKPNKSTPVNEENKKEYVRLYNWFLRGI 360

595 EAQFLALQKGFNEVPOHLLKTFDEKELELIIICGLGKIDVNDKVNTRLKHCTPDSNIVK 654
 361 EAQFLALQKGFNEVPOHLLKTFDEKELELIIICGLGKIDVNDKVNTRLKHCTPDSNIVK 420

655 WFKAVEFDEERARLLQFVTGSSRVLQGGKALQGAAPRFTTHQDACTNNLPKHAH 714
 421 WFKAVEFDEERARLLQFVTGSSRVLQGGKALQGAAPRFTTHQDACTNNLPKHAH 480

715 TCFNRIDIPPEYSEKYLEKLLTAIETCGFAVE 748
 481 TCFNRIDIPPEYSEKYLEKLLTAIETCGFAVE 514

UT 8
 61120
 ABB61120 standard; Protein; 1035 AA.
 ABB61120;
 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 10152.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS WO200171042-A2.
 PN 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL05223.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 10152; 21pp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB357737-AB372072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1035 AA;

Query Match 53.8%; Score 2173; DB 22; Length 1035;
 Best Local Similarity 43.2%; Pred. No. 1.4e-195;
 Matches 456; Conservative 104; Mismatches 154; Indels 342; Gaps 16;

Qy 7 RRRGPKVRLTLVLCARNLVKDFPRLPDPFAKVVDGSGQCHSTDTVKNLDPKWNQHYD 66
 Db 8 RRRGTHKVRITILCARNLARKDLFELPDPFAKVVDGSGQCHSTDTVKNLDPKWNQHYD 67

Qy 67 LYTKSDSVTISVWVHKHKKQAGFLGCVLRLSNAINRLKDTGYQRLDCKLGPNDND 126
 Db 68 LFLIGDGAITITVWNRKIHK--GSGFLGCVRIAPAFNIQSLKAGAGFQLDLGLKLSFPDDE 125

Qy 127 TVRQGVWSLQGRDR-----ITGGVQVVDCSR--LFDNDLPDGWEERTASGRIOY 175
 Db 126 LVRGQIIISLLSKDGPSSGNPLAIVGSDVGRGPEDDSDSLPEGWERTTNGRVVY 185

Qy 176 LNHITTTQWERTP-----
 Db 186 VNHATKSTQWDRPQGVGVSSGNPLAIVGSDVGRGPEDDSDSLPEGWERTTNGRVVY 191

Qy 192 -----ASEYSSSPGRPLSCFV-----
 Db 246 HRSRDLVTSADERRHSTELSSVGKENTSPPTVYSATTPGKKTSSSSSSAGGRTLEQ 245

Qy 207 -----DENTPISGT-----N 216
 Db 306 RPTNEPATSTSTTSASVRLHSDNHNKTPKHQTNHAPPSTPTSPGQOYVNGNAQN 365

Qy 217 GAT-----CGQSSDPRLAER-----
 Db 231

356 GSTSGSGQAQPOSASNGWTQEDATTTSPPTTTPPHRSQSPPTPNISPPASVTPSA 425
232 -----RVRSQRHENYMSRTHLHTPP----- 251
426 NGNVHSPNANSTPAGSGGSRSYTAATPGQSRSSRQOGBESSRRSSRGRNGGTS 485
252 -----DLPEGVEQRTTQGGVFLHTCTGVTWHDPR 283
486 GGGGGGSGGQYASAAATAANQAARPFDDPPGYEMRTTQGGVYFIHTPTGVTWHDPR 545
284 VPRDL--SNINCEELGFLPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRQ 341
546 IPRDPDTQHLTDAIGLPLSGWEQRTASGRVYFVDHNNRTTQTDPRLSGSLQIRRG 605
342 N-----QLKDOQQQQVSL-----CPDTE 361
606 TVPPTSAANATPAPPATPATPSAAAVFPQATPSANATPTLTTPPHRIVPDLFQ 665
362 CL-----TVPRYKDLVQKILRQELSQOQOAGHCRIEVSREEIFESYRQVMQRPK 416
666 GLEGADILPKYRRDLVGKURALATELCTWQPQSGHCRLEVSRENEIFESSYRLIMQRAK 725
417 DLWKLMTKFBEGELDYGGVAREWLVLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVN 476
726 DMRLKMYKFBEGELDYGGVAREWLVLLSHEMLNPPYGLFOYSRDDIYTLQINPDSGVN 785
477 PEHLSYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDMELVDPDLHNSLWMI 536
786 PDHLSYFHF-----LNLKPIITGDIQGVDPDLHRSLTWM 819
537 LENDITGVLDHTFCVEHNAYGEIIQHELKPKNGKSIPIVNEENKKEYVRLVYNNWPLRGIEA 596
820 LESNTSGIESTFVSNNFSGALVHSLKPGGASIPVTEENKKEYVRLVYNNWPLRGIEQ 879
597 QFALQKGFNEVIPHLLKTKTDEKELELIGLGKIDVNDKYNTRLKCHCTPDSNVKWF 656
880 QFALQKGFCELIPLSHLRPFDERELELVIGGSIIDVNDWRNTRLKCHCTNEITQVLWF 939
657 WKAVEFFDEERARLLQVTSRVPVLOGFKALO--GAAGPLFTIH-QIDACTNNLPK 712
940 QWVESYSEMRARLLQVTSRVPVLOGFKALQSGTGAAGPLFTIHLTADVPTQNLPK 999
713 AHTCFNRIDIPYSEYKELKLTATEETCGFAVE 748
1000 AHTCFNRIDIPYETQYLLCDKLTQAVEETCGFAVE 1035

!ULT 9
!13385

AAW13385 standard; Protein; 766 AA.

AAW13385;

10-JUL-1997 (first entry)

Human protein ubiquitin ligase pub2.

Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase;
p53; cell cycle; transgenic animal.

Homo sapiens.

WO9712962-A1.

10-APR-1997.

04-OCT-1996; 96WO-US15930.

04-OCT-1995; 95US-0539205.

(COLD-) COLD SPRING HARBOR LAB.

Beach D, Caligiuri M, Nefsky B;

XX WPI; 1997-226206/20.
DR N-PSDB; AAT47041.
XX Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved
PT in cdc25 phosphatase and p53 ubiquitination, and regulate cell
PT growth and proliferation
XX Claim 1; Page 80-84; 108pp; English.
XX Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and
CC pub3 (AAW13386) are homologues of fission yeast pub1 (AAW13387) and
CC were identified from cDNA clones (AAT47040-42) obt'd. e.g. from a
CC keratinocyte cDNA library. Pub polypeptides can be produced in
CC transfect host cells. They can control the steady state level of
CC cdc25 phosphatase, the degree of CDK kinase (e.g. cdc2)
CC dephosphorylation and the steady state level of p53 (controlling
CC the degree of cell cycle regulation of p53). They can regulate
CC cell or tissue differentiation, or cell growth or proliferation by
CC affecting other proteins, can be a specific (ant)agonist of wild-
CC type protein function and may be used as immunogens to elicit a
CC specific immune residue.
XX Sequence 766 AA;
SQ
Query Match 38.3%; Score 1545; DB 18; Length 766;
Best Local Similarity 40.8%; Pred. No. 2.1e-136;
Matches 336; Conservative 121; Mismatches 234; Indels 132; Gaps 18;
QY 1 MSNPRRRNGPVKRLITVLCANLKVQDFRLLPDPFAKVVDGSGQCHSTDVXKTLDPK 60
Db 1 MSNSAQSR-----RIRVIAADGLYKRDVFRFPDPAVLTVDGE-QTHTTTAKKTLPY 55
QY 61 WNOHYDLIGKSDSVTISVNMKKHKQAGFLGCVRLLSNAINRLKDTGYORL--DLC 118
Db 56 WNETFVNVTDNSTIAIQVFDQKFP-KKKGQSGFLGVINLRVGDVLDLAIGGDEMLICDLK 114
QY 119 KLGNNDNDTVRQIVVSLQ----- 137
Db 115 KL--NENTVVHGKIIINLSTTAQLTLQVPSSAASGARTORTITNDPOSSKSSVSNNPA 172
QY 138 -----SRDRIGTGQGVDCSRLPDN-----DLPDGWEERTASGRIOYLNHTRTTO 184
Db 173 SSRAGSTRONAPAASPASEPRTFSFEDQVGRLLPGWERTDNLGRYYVDHNRSTT 232
QY 185 WERTP-----RPASEYSPGRPLSCFVDENTPIGCTNGATCGQSDPPLAERRVRSORH 238
Db 233 WIRENLSSVAGAAAAAELHSSAS--SANVTGQVOPSSNAA-----RRTEASVLT--- 279
QY 239 RNYMSRTHLHTPPDLPEGVEQRTTQGGVFLHTCTGVTWHDPRVPRDLSNIN----- 292
Db 280 ----SNATTAGSGELLFGWEQRTTPGGRFYFVDHNRTRITTVWDRRQRTIRSYGGPNAT 335
QY 293 -----CBELGFLPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRQNLKDO 347
Db 336 IQQQPVSQLGFLPSGWEMLTNTARVYFVDHNTKTTTWDPLPSSL-----DQ 384
QY 348 QQQQVSLCPDDETECLTVPRYKDLVQKILRQELSQO--QPOAGHCRIEVSREEIFEE 405
Db 385 -----NVPQYKDFRKLIFP-----LSQPALPLFPQCHIKVRRNHIFED 425
QY 406 SYRQVMQRPKDLWKLMTKFBEGELDYGGVAREWLVLLSHEMLNPPYGLFOYSRDDIY 465
Db 426 SYAEMRQSATDLKRLMIKFDGEDGLDYGGLSREYFFLLSHEMFNPFYCLFEYSVDNY 485
QY 466 TLQINPDSAVNPEHLSYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDMELV 525
Db 486 TLQINPHSGINPEHLNAYFKPIGRVIGLAIFHRFVDAFFVVSFYKMLQKKVTLQDMESM 545
QY 526 DPDLHNSLVMLENDITGVLDHTFCVEHNAYGEIIQHELKPKNGKSIPIVNEENKKEYVRLY 585
Db 546 DAETYSRLVILNDITGVLDLTFSVEDNCFBGEVVTIDLKPNGRNIEVTEENKRYVDLV 605

BLT 11	QY	11	PVKLRLLTVLCANLVKKDRFR-----LDPFAKVVVDGSGOCH 48
6104	DB	15	PASSRLCRVIAADGLYXRDRFRKSYILAILLVRLITSGAFAGFPDFEAVATVGGE-QTH 73
ABU26104;	QY	49	STDTVQNTLDPKMNQHYDLYIGKSDSVTISWNHKKIHKQAGFLGCVRL-- 100
16-APR-2003 (first entry)	DB	74	TTSVIKKTLPYWNEMFDRVNEBSILAIQIFDQKXF-KKQDQFLGVINVRIGDVIDLQ 132
Aspergillus fumigatus essential gene protein #762.	QY	101	-----SNAINRLK-----DTGQRIIDLCKLPNDNDTVRGQIVSL----- 136
Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection; cancer; contamination; biofilm; antibody; immune response.	DB	133	MGDGESLPTRHSCDVRSLRFLRLLEMLEMLTDLK--SNDLVVHGKLIINLSTNLS 190
Aspergillus fumigatus.	QY	137	-----QSRDRIGTGQVVDCS----- 152
WO200286090-A2.	DB	191	PNTQANGLHSHVQSSTSSGLVPQVAPSSSHPAASGTAPVDPASNPSPINPQVPSSTR 250
31-OCT-2002.	QY	153	-----RUFND-----LDPWEERTASGRIOYLNHITRTTQ 184
23-APR-2002; 2002WO-US13142.	DB	251	PSSTAAPASAAAGAAVNSHGSRTNLSFSDSQGLPAGWERREDNLGRTYVDHNRTRTT 310
23-APR-2001; 2001US-285697P.	QY	185	WERPTRPASEYSSPGRPLSCFVDENTPISGTNGATCGQSDPRLAERRVRSORHNYM-- 242
27-APR-2001; 2001US-287066P.	DB	311	W---TRPSSNYNEHAQ-----RSQREANWQLERRAQAQSRMLP 344
05-JUN-2001; 2001US-295890P.	QY	243	-----SRTHLHTPP-----DLPEGYEQRTTQOQVY 268
09-JUL-2001; 2001US-303899P.	DB	345	EDRTGANSPLPSSQAHTPPAGGGSANAVSMMATGATTAGTGELPGWEGQRTTPREGPY 404
31-AUG-2001; 2001US-316362P.	QY	269	FLHTQTVSTWHDPRVPRDI-----SNIN-----CEELGPLPGWEIRNTATGRVY 314
(BLIT-) ELITRA PHARM INC.	DB	405	FVDENRTTTTWDPRQRYIMYGQNGANGNTTIQQQPVSQLGFLPSGWEMLTNTARY 464
Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;	QY	315	FVDHNNRTTQTPRLSANLHLVLRNQNLKQDQOQQVSLCPDDTECLTVPRYKRDVLQ 374
WPI; 2003-093124/08.	DB	465	FVDHNTKTTTWDPRLPSSL-----DQ-----GVPOYKXDFR 497
New purified or isolated nucleic acids of essential genes of Aspergillus fumigatus, useful for treating or preventing infections by A. fumigatus, or for treating a non-infectious disease in a subject e.g. cancer	QY	375	KLKILRQELSQQP-----QAGHCRIEVSREIEESYRVQVMKRPKDLKELMKFERGE 430
Disclosure; Page : 175pp; English.	DB	498	KLIVFR-----SQPALRMGQCHVKVRNNIFEDSYAEIMROSADLKKRLMKFDGSD 552
The invention relates to novel purified or isolated nucleic acids of essential genes of Aspergillus fumigatus. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as A. fumigatus, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by A. fumigatus, or to prevent or inhibit formation on a surface of a biofilm comprising A. fumigatus. The polynucleotides are useful for expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of A. fumigatus to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, for raising anti-protein antibodies, as an antigen to raise anti-DNA antibodies or to elicit another immune response, and for identifying polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction. The polypeptides may be used to raise antibodies or to elicit immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case of virulence factors. This sequence represents a protein of one of the essential genes of Aspergillus fumigatus of the invention.	QY	431	GLDYCGVARELYLLSHEMLNPYVGLFOYSRDDIYTLQINPDASAVNPEHLSVFFHVGRI 490
	DB	553	GLDYGLSGEFFFLLSHEMFNPFVCLFEYSADHNYTLQINPHSGVNPHEHLYFKFGRVV 612
	QY	491	GNVPHGYIDGGTLPFYKQLGKSLTLDMLVDPDLHNSLVWLENDITGVLDHTFC 550
	DB	613	GLAIFHRFLDFFIGAFYKMLRKKVSLQDMGVEDLHRLNTWTWMDIEGVLELTF 672
	QY	551	VEHNAYGRIIQHELPNGKSGIPVNEENKKEYVLYVNRFLRGIEAQFLAQKGFNEVIP 610
	DB	673	VDEKFGERRRIDLKPGGRDIPVTNENKAEYVLYVTEWKIVKEVEQFNAMFGFNELIP 732
	QY	611	QHLKTFDEKELELIICGLKIDVNDKVNTRLKHCTPPDSNIVKFWKAVEFPDEBRRAR 670
	DB	733	ADLVNFDERELELLIGIADIDVDDWKKHTDYRGYQESDEVIONFVKIVRSWDABQSKR 792
	QY	671	LLQFTGSSRVPLOGKALQAGAPRLFTTHQIDACTNMLPKAHTCFNRIDIPPYESYEK 730
	DB	793	LLQFTGTSRIPVNGFKDLOGSGPRFTIEK-SGDPALPKSHTGFRDLDPYKSYET 851
	QY	731	LYEKLLTAIESTCGFAVE 748
	DB	852	LEHKMSIAVESTLFGQE 869
Sequence 869 AA;			
ery Match	XX		
st Local Similarity	XX		
atches 332; Conservative 114; Mismatches 229; Indels 243; Gaps 21;	DT		

RESULT 12

AAW93167

ID AAW93167 standard; Protein; 975 AA.

XX

AC AAW93167;

XX

DT 24-MAY-1999 (first entry)

XX

FT	Domain	618..724
----	--------	----------

FT	Domain	618..724
----	--------	----------

Disclosure; Page 68-71; 79pp; English.

This invention describes a novel assay identifying an inhibitor of ubiquitin-mediated proteolysis of a IkappaB polypeptide. The method comprises: (a) providing a ubiquitin-conjugating system including the IkappaB polypeptide and a HECT (Homologous to E6-Ap Carboxyl Terminus) ligase and ubiquitin, to promote ubiquitination of the IkappaB polypeptide by the HECT ligase; (b) contacting the ubiquitin-conjugating system with a candidate agent; (c) measuring a level of ubiquitination of the polypeptide in the presence of the candidate agent and (d) comparing the measured level of ubiquitination in the presence of the candidate agent with ubiquitination of the IkappaB polypeptide in the absence of the candidate agent; where a statistically significant decrease in ubiquitination of the IkappaB polypeptide in the presence of the candidate agent is indicative of an inhibitor of ubiquitination of the IkappaB polypeptide. The assays can be used to identify compounds which modulate binding and/or ubiquitination of an IkappaB (or other cellular or viral substrate) by a HECT ligase, such as RSC or KIAAN. Such modulators can be used e.g. in the treatment of proliferative and/or differentiative disorders, to modulate apoptosis, in the treatment of viral infections, and in the treatment of tissue wasting disorders e.g. cachexia secondary to infection or malignancy, cachexia secondary to human AIDS, inflammatory diseases, parasitic diseases, tuberculosis and high dose IL-2 therapy; rheumatoid arthritis, rheumatoid spondylitis, osteoarthritis, gouty arthritis and other arthritic conditions, sepsis, respiratory distress syndrome, cerebral malaria, chronic pulmonary inflammatory disease, silicosis, pulmonary sarcoidosis, bone resorption diseases, reperfusion injury, graft versus host reaction, allograft rejections, Crohn's disease, ulcerative colitis, or pyresis, in addition to a number of autoimmune diseases such as multiple sclerosis, autoimmune diabetes, systemic lupus erythematosus; and ENL in leprosy, HIV, and AIDS. This sequence represents a human KIAAN ligase which is used in the method of the invention.

Sequence 927 AA;

ary Match 33.5%; Score 1354.5; DB 20; Length 927;
 at Local Similarity 34.0%; Pred. No. 3e-118;
 tches 329; Conservative 123; Mismatches 232; Indels 285; Gaps 28;
 4 PGRNRGPNV-----KRLTLVLCARNLVKKDFRLLPDPFAK 38
 13 PGRGRARPCWKRSEMATCAVEVFGLLEDENSRIVRVRIAGIGLAKKDILGASDPYVR 72
 39 VV-----VDSGGOCHSTDTVKTLPKXN-----QH-----YDLVIGK 71
 73 VTLDPMMGVLTSTVQTKTKSLNPKWNEELFRVHPQOHRLLFEVDENRLRDPDGLQ 132
 72 SD--SVTISVMN-----HKIKHKQAGLGCVRLLSNAINRLKDTGYQRL 115
 133 VDVPLPLPTENPLRERYTKDFVLPKSHKSRVKGVL-----RLKWT----- 176
 116 DLCKLGNPDNTRVGQIVVLSQRDRITGQGVVD-----C---SRLFNDLDPGHEER 167
 177 YLPKTSGSDD-----NAQAELEFGWVLDQPDAAACHLQQQOQSPPLPGWEERQ 228
 168 TASGRIOYLNHLTRITQWERPT-----RPASE----- 194
 229 DILGRYYVNHESRTQWKRPTQDNLTDAENGNILOQAQPAFTTRQISEETESVDNQE 288
 195 -----YSS---PGRFLSCFVDNTPISGT-----NGATCGQSSDPR 227
 289 SSENWEIIEDEATWYSQAQPPSPSSNLDVPHLAEELNARLTIFGNSAVSQPASSN 348
 228 LAERR-----VRSQRHNYM-----SRTHLHTPPD- 252
 349 HSSRRGSLQAYTFESOPTLPLVLLPTSSGLPPGWEEKQDERGSRVYVDHNRITWTPTV 408
 253 -----LPEGYEORTTQQQGVYFLH 271
 409 QATVETSQLTSSQSSAGPQSQASTSDSGQVQTPQSEIEQGFPLPKGWVHRAPNRPFFID 468
 272 TQGVSTWHDPR--VPR-----DLNINCEELGPLPPGWERTNATGRVYFVDHNR 321

Db 469 HNTKTTWEDFLKXPAHLRGKTSLDTSN-----DLGPLPGWEERTHTDGRIFVINENIK 524
 Qy 322 TQQTDPRLSANHLVLRQNLQDQOQQQVSLCPDDTECLTVPR--YKEDLVQKLKIL 379
 Db 525 RTQWEDPRL-----ENVAITGPAVPYGRDYKRYEYFF 556
 Qy 380 ROELSQQQPAQCHRIEVSREBEIPESYRQVMKRPKDLWK-RLMIXFRGEEGLDYGGVA 438
 Db 557 RRLKKQNDIPNKFEMKLRATVLEDSYRIMGVKRADPLKARLWIEFDGEEKLDYGGVA 616
 Qy 439 REWLYLSHMLNPPYGLFQYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMGMAVPHG 497
 Db 617 REWFFLISKEMFNYYGLFEYSATDNYTLQINPNSGLCNEHDLSYFKFPGVAGVAVYHG 676
 Qy 498 HYIDGGFTLPFFYKOLLGKSITLDDMELVDPDLHNSLVWILENDITGVLDHTFCVHNAYG 557
 Db 677 KLIDGFFIFFYMWMLHLPITLHDSVDSVYNSLRWILENDPT-ELDLAFIIDEELFG 735
 Qy 558 EITQHELKNGKSIPIVNEENKKEYVRLVYNNRFLRGIEAQFLAQKQNEVPIPOHLLKTF 617
 Db 736 QTHQHELKNGSGSEIVVTKNKKKEYILVQWRFNRIQKMAAFKEGFEFELIPQDLIKIF 795
 Qy 618 DEKELELIICGLGKIDVNDKYNTRIKH-CTPDSNIVKWFKAVEFFDEERRARLLQFVT 676
 Db 796 DENELELLMCGUGVDVNDWRHRTKFKNGYSANHQVIOFWKAVLWMDSEKRIILLQFVT 855
 Qy 677 GSSRVPLQGFALQGAAGPRLFTIHQIDACTNNLPKANTCNRIDIPYSEYKLYEKL 736
 Db 856 GTSRVPWNGFAELYGSGNGPSQFTVEQ-GTPEKLPRAHTCFNRLDLPYSEFEELMDKLQ 914
 Qy 737 TAIEETCGF 745
 Db 915 MAIENTQGF 923
 RESULT 15
 AA32725
 ID AA32725 standard; Protein; 927 AA.
 XX
 AC AA32725;
 XX
 DT 24-MAR-2003 (first entry)
 XX
 DE Human NEDD4 long form protein.
 XX
 KW Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis; poliomyelitis; HIV; measles; protein therapy; human; NEDD4.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 381..406
 FT notes="WW domain"
 FT 821..923
 FT notes="HECT domain"
 XX
 PN WO200290549-A2.
 XX
 PD 14-NOV-2002.
 XX
 PF 12-MAR-2002; 2002WO-IB02106.
 XX
 PR 12-MAR-2001; 2001US-275224P.
 PR 31-JUL-2001; 2001US-308958P.
 PR 07-DEC-2001; 2001US-340170P.
 XX
 PA (PROT-) PROTOLOGICS LTD.
 XX
 PI Greener T, Moskowitz H, Reiss Y, Alroy I;
 XX WPI; 2003-111976/10.
 DR N-PSDB; AAD50459.

New protein complex comprising HECT-RCCL1, viral maturation scaffolding protein (VNSP), and/or HIV gag protein, useful for treating viral infections, such as lymphosarcoma, HIV, hepatitis, poliomyelitis, measles, or Ebola -

Disclosure; Fig 15; 150pp; English.

The invention relates to a method for modulation of viral maturation. The invention also provides an isolated protein complex comprising a HECT-RCCL1 selected from HECT-NW, HECT-RCCL1, Gag protein, Gag late domain, P13, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, GPase, E2 enzyme, teg101, cullin, HEC1, HEC2, HEC3, Nedd4-like protein or clathrin. The complexes, proteins, antibodies and methods are useful for treating viral infections, such as lymphosarcoma, human immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola and for inhibiting budding in a subject. They are also useful in diagnostic assays for determining whether a cell is infected with a virus and for characterizing the nature, progression and/or infectivity of the infection. The invention is also useful in protein therapy. The present sequence is human NEDB4 protein used to illustrate the method of the invention.

Sequence 927 AA;

ery Match 33.5%; Score 1354.5; DB 24; Length 927;
st Local Similarity 34.0%; Pred. NO. 3e-118;
tches 329; Conservative 123; Mismatches 232; Indels 285; Gaps 28;
4 PGRNRGPNV-----KLRITLVCAKNLVKDFRLLPDPFAK 38
13 PGRGRAPVCKRSEMATCAVEVGGLEDENSRIVRVVIAGIGLAKKILGASDPYVR 72
39 VV-----VDGGGCHSTDTVKNTLPKMN-----QH-----YDLYIGK 71
73 VLYDPMNGVLTSTVQTKTIKSLNPKWNEELLFRVHQHRLLEFVDFENRLTRDDFLGQ 132
72 SD--SVTISVNV-----HKIKHKQAGFLGCVRLLSNAINELKDTGYQRL 115
133 VDVPYLPPTENPLRLERYPTKDFVLHPRSHKSVKGYL-----RLKWT----- 176
116 DLCKLGPNDNDTVRQQIVVSLQSDRIDGTGGQVVD-----C-----SRIFDNDLPDGEERR 167
177 YLPKTSGEDD-----NAEQAEELPGVGVLDQPDAAHQQLQQQEPSPPLPGWEERQ 228
168 TASRIQVNLNHTRTQWERT-----RPASE----- 194
229 DILGRYYVNHESRETOKRPTPDNLDAENGNIQLQAQRAFTTRQISEETESVDNQE 288
195 -----YSS-----PGRPLSCFVDENTPISGT-----NGATCGQSSDPR 227
289 SSENWEIREDATWYSSQAFSPSPSSNLDVPHLAEELNARLTIFGNSAVSQPASSN 348
228 LAER-----VRSQRHNYM-----SRTHLHTPPD- 252
349 HSSRRGSLQAYTFEEQPTLPVLLPTSSGLPPGWEKQDERGRSYVDHNSRTTWTKPTV 408
253 -----LPEGYEORTTQGGQVVELH 271
409 QATVTSQITSSQSAGQSASTSDSQQTQSEIEQGFPLKGVNEVRHAPNGRFFPID 468
272 TGTGVSTWHDPR--VPR-----DLNINCEBELGPLPGWEIRNTATGRVYFVDHNR 321
469 HNTKTTTWTEDPLKIPAHURGKTSLDTSN-----DLGFLPPGWEERTHTDGRIFVNHNIK 524
322 TTQFTDPLRLSANHLVLRNQKLDQQQQVVSICPDDTECLTVPR--YKDLVQKLKIL 379
525 RTQWEDPRL-----ENVAITGPAVYGRDYRKRYEFF 556
380 ROELSQQQPAQCHRIEVSREBEIFESYRQVQMKRPKDLWK-RLMIXFRGEEGLDYGVGA 438
557 RRLKKQNDIPNKFEMKLRRATVLEDSYRRINGVKRADFLKARLWIEFDGEGKLDYGVGA 616

QY 439 REMLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFYGRIMGMAVEHG 497
Db 617 REWFLLSKEMFPYGLFYSATDNYTLQINPNSGLCNEDHLSYFKFGRVAGNAVYHG 676
QY 498 HYIDGGTLPFYKQLIGKSTLDDMELVDPDLHNSLVWILENDITGVLDHTFCVBNHNAV 557
Db 677 KLDDGFFIRPFPYKMLHKKPITLHDMESVDSEYNSLRWILENDPT-ELDLRFIDDEELFG 735
QY 558 BIIQHEIKPKNGKSPVNEENKKEYVRLVYNWRFLGIEAQFLALOKGNEVIPQHLKTF 617
Db 736 QTHQHELANGSEILVTNNKKEYIYLVQWRFVNRIOQMAAFKEGFFELIPQDLIKIF 795
QY 618 DEKELELIICGLGKIDVNDKVNTRLKH-CTPDSNIVKFWKAVREFFDEERARLLQFVT 676
Db 796 DENELELLMCGLGVDVNDWREHTKYKNGYSANHQVQWFKAVLMDSEKIRLLIQFVT 855
QY 677 GSSRVPLQGFKALQGAAGPELFTIHQIDACTNNLKAHTCENRIDIPPYESYEKLYEKL 736
Db 856 GTSRVPNGFAELYGSGNGSQSFVTEQW-GTPEKLPRAHTCFNRLLDLPPIESFEELWDLQ 914
QY 737 TAIBETCGF 745
Db 915 MAIENTQGF 923

Search completed: February 20, 2004, 15:28:33
Job time : 45.6968 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
protein - protein search, using sw model
n on: February 20, 2004, 15:26:11 ; Search time 15.7634 Seconds
(without alignments)
2007.718 Million cell updates/sec
t1e: US-10-009-945-4
xfect score: 4038
quence: 1 MSNPGRRNGPVKRLTLVLC.....EKLYEKLTAIEETCGPAVE 748
oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
arched: 328717 seqs, 42310858 residues
tal number of hits satisfying chosen parameters: 328717
nimum DB seq length: 0
ximum DB seq length: 2000000000
st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
tabase :
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pdp:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pdp:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pdp:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pdp:*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pdp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match %	Length	DB ID	Description
1	3939.5	97.6	735	3	US-08-539-205A-2
2	3939.5	97.6	735	4	US-09-392-163A-2
3	1564	38.7	766	3	US-08-539-205A-4
4	1564	38.7	766	4	US-09-392-163A-4
5	1354.5	33.5	927	3	US-08-895-601-6
6	1353	33.5	834	3	US-08-539-205A-6
7	1353	33.5	834	4	US-09-392-163A-6
8	1336	33.1	854	2	US-09-070-060-4
9	1336	33.1	854	3	US-09-357-746-4
10	1329	32.9	852	2	US-09-070-060-3
11	1329	32.9	852	3	US-09-357-746-3
12	1298.5	32.2	906	3	US-08-630-916A-48
13	1134.5	28.1	683	3	US-08-630-916A-46
14	534.5	13.2	874	2	US-08-247-904B-8
15	534.5	13.2	874	3	US-08-767-942A-21
16	513	12.7	866	3	US-08-100-692-1
17	513	12.7	866	2	US-08-674-030-1
18	450.5	11.2	1083	3	US-08-895-601-5
19	219	5.4	486	3	US-08-348-518C-2
20	207	5.1	472	3	US-08-348-518C-5
21	207	5.1	472	3	US-08-476-509B-5
22	162.5	4.0	54	3	US-08-630-916A-124
23	158	3.9	448	3	US-08-476-509B-2
24	156.5	3.9	55	3	US-08-630-916A-75
25	154.5	3.8	454	3	US-08-348-518C-4
26	154.5	3.8	454	3	US-08-476-509B-4
27	153	3.8	224	3	US-08-630-916A-50

Sequence 117, Appl
Sequence 74, Appl
Sequence 73, Appl
Sequence 84, Appl
Sequence 118, Appl
Sequence 25, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 36, Appl
Sequence 116, Appl
Sequence 32, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 115, Appl
Sequence 24, Appl
Sequence 15, Appl
Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-539-205A-2
; Sequence 2, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-539-205A-2

Query Match 97.6%; Score 3939.5; DB 3; Length 735;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;
QY 1 MSNPGRRNGPVKRLTLVLCANLKKDFRLLPDPFAKVVVDGSGGCHSTDVXKNTLDPK 60
Db 1 MSNPGRRNGPVKRLTLT-----GLDPFAKVVVDGSGGCHSTDVXKNTLDPK 47
QY 61 WQHYDIYIGKSDSVIISVWVNHKKIKQAGFLGCVRLLSNAINRLKDTGYORLDLCKL 120

48 WNOHYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL 107
 121 GNDNDYRGQIWSLSQSDRICTGGQVWDCSRLPFDNDLPDGEERTASGRIOYLNHIT 180
 108 GNDNDYRGQIWSLSQSDRICTGGQVWDCSRLPFDNDLPDGEERTASGRIOYLNHIT 167
 181 RTQWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGSSDPRLAERRVRSQRHN 240
 168 RTQWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGSSDPRLAERRVRSQRHN 227
 241 YMSRTHLTPDLPGEYQORTTQGGVYFLHTQTVSTWHDPRVPRDLNINCEELGPLP 300
 228 YMSRTHLTPDLPGEYQORTTQGGVYFLHTQTVSTWHDPRVPRDLNINCEELGPLP 287
 301 PGWEIRNTATGRVYFDHNNRTTQFTDRLSANLHLVLRNQNLKDDQGGVYSLCPDDT 360
 288 PGWEIRNTATGRVYFDHNNRTTQFTDRLSANLHLVLRNQNLKDDQGGVYSLCPDDT 347
 361 ECLTVPRYKRDVQKILRQELSQQPGAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
 348 ECLTVPRYKRDVQKILRQELSQQPGAGHCRIEVSREEIFEESYRQVMKRPKDLWK 407
 421 RLMIKFRGEGLDYGGVAREMILYLLSHEMLNPPYGLFOYSRDDIYTLQINPDNAVPEHL 480
 408 RLMIKFRGEGLDYGGVAREMILYLLSHEMLNPPYGLFOYSRDDIYTLQINPDNAVPEHL 467
 481 SYFHFVGRIMGAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLNSLVWILEND 540
 468 SYFHFVGRIMGAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLNSLVWILEND 527
 541 ITGVLDHTFCVEHNAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRLRGIEAQFLA 600
 528 ITGVLDHTFCVEHNAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRLRGIEAQFLA 587
 601 LQKGFNEVIFQHLKTFDKELELIICGLGKIDVNDKVNTRLKCHCTPDSNIVKWFKAV 660
 588 LQKGFNEVIFQHLKTFDKELELIICGLGKIDVNDKVNTRLKCHCTPDSNIVKWFKAV 647
 661 EFPDEERRARLLOFTVGSSRVPLQGFKALQGAAPRLFTTHQIDACTNNLPKHAFTCFNRI 720
 648 EFPDEERRARLLOFTVGSSRVPLQGFKALQGAAPRLFTTHQIDACTNNLPKHAFTCFNRI 707
 721 DIPPEYSEKLYEKLTAETTCGFAVE 748
 708 DIPPEYSEKLYEKLTAETTCGFAVE 735

MULT 2
 .09-392-163A-2
 Sequence 2, Application US/09392163A
 Patent No. 6503742
 GENERAL INFORMATION:

APPLICANT: Beach, David H.
 APPLICANT: Caligiuri, Maureen
 APPLICANT: Neisky, Bradley
 TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
 NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA

COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/392,163A
 FILING DATE:

PRIOR APPLICATION DATA:
 FILING DATE:

APPLICATION NUMBER: US 08/539,205
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: CSV-005.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 832-1000
 TELEFAX: (617) 832-7000
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 735 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-392-163A-2

Query Match 97.6%; Score 3939.5; DB 4; Length 735;
 Best Local Similarity 98.0%; Pred. No. 0;
 Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;

QY 1 MSNFGRRNGPVLRLTVLCAKLVKKDFRLLPDPFAKVVVDGSGGCHSTDTVKNTLPK 60
 DB 1 MSNFGRRNGPVLRLTVLCAKLVKKDFRLLPDPFAKVVVDGSGGCHSTDTVKNTLPK 47
 QY 61 WNOHYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL 120
 DB 48 WNOHYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL 107
 QY 121 GPNNDYVRGQIWSLSQSDRICTGGQVWDCSRLPFDNDLPDGEERTASGRIOYLNHIT 180
 DB 108 GPNNDYVRGQIWSLSQSDRICTGGQVWDCSRLPFDNDLPDGEERTASGRIOYLNHIT 167
 QY 181 RTQWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGSSDPRLAERRVRSQRHN 240
 DB 168 RTQWERPTPASEYSSPGRLSCFVDENTPISGTNGATCGSSDPRLAERRVRSQRHN 227
 QY 241 YMSRTHLTPDLPGEYQORTTQGGVYFLHTQTVSTWHDPRVPRDLNINCEELGPLP 300
 DB 228 YMSRTHLTPDLPGEYQORTTQGGVYFLHTQTVSTWHDPRVPRDLNINCEELGPLP 287
 QY 301 PGWEIRNTATGRVYFDHNNRTTQFTDRLSANLHLVLRNQNLKDDQGGVYSLCPDDT 360
 DB 288 PGWEIRNTATGRVYFDHNNRTTQFTDRLSANLHLVLRNQNLKDDQGGVYSLCPDDT 347
 QY 361 ECLTVPRYKRDVQKILRQELSQQPGAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
 DB 348 ECLTVPRYKRDVQKILRQELSQQPGAGHCRIEVSREEIFEESYRQVMKRPKDLWK 407
 QY 421 RLMIKFRGEGLDYGGVAREMILYLLSHEMLNPPYGLFOYSRDDIYTLQINPDNAVPEHL 480
 DB 408 RLMIKFRGEGLDYGGVAREMILYLLSHEMLNPPYGLFOYSRDDIYTLQINPDNAVPEHL 467
 QY 481 SYFHFVGRIMGAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLNSLVWILEND 540
 DB 468 SYFHFVGRIMGAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLNSLVWILEND 527
 QY 541 ITGVLDHTFCVEHNAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRLRGIEAQFLA 600
 DB 528 ITGVLDHTFCVEHNAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRLRGIEAQFLA 587
 QY 601 LQKGFNEVIFQHLKTFDKELELIICGLGKIDVNDKVNTRLKCHCTPDSNIVKWFKAV 660
 DB 588 LQKGFNEVIFQHLKTFDKELELIICGLGKIDVNDKVNTRLKCHCTPDSNIVKWFKAV 647
 QY 661 EFPDEERRARLLOFTVGSSRVPLQGFKALQGAAPRLFTTHQIDACTNNLPKHAFTCFNRI 720
 DB 648 EFPDEERRARLLOFTVGSSRVPLQGFKALQGAAPRLFTTHQIDACTNNLPKHAFTCFNRI 707
 QY 721 DIPPEYSEKLYEKLTAETTCGFAVE 748
 DB 708 DIPPEYSEKLYEKLTAETTCGFAVE 735

348 QQQVVSLCPDDTECLTPRYKRDIVQKILROELSQQ--QPOAGHCRIEVSREEIFEE 405
385 -----NVEQYKRDFFRRKLIYF--LSQALHPLPGQCHIKVRNHFED 425
406 SYROVMXRPKDLWKRLMTKRGEGGLDYGVAREWLYLLSHENLNPYYGLFOYSRDDIY 465
426 SYAEMRQSATDLKKRLMIKFDGEDGLYGLSREYFFLLSHENMFNYCLFEYSVDNY 485
466 TQIINPDSAVNPEHLSYFHFVGRINGMAVPHGHYIDGGFTLPFFYKOLLGKSITLDDMELV 525
486 TQIINPHSGINPEHLNYPKFIQVIGLAIPIHRRFVDAFFVVSFYKMLQKVTLDQVSM 545
526 DPDLNLSVWILENDITGVLDHTFCVENAYCEIIOHELKPNKXSIPVNEENKEIYRLY 585
546 DAERYKSLWILENDITGVLDHTFVEDNCEGVVTIDLPKNGRNIETENKGEYVDLV 605
586 VNWRELRGIEAQLALQKGFNEVPIQHLTKTFDEKELELIICGLGKIDVNDWKVNTRLKH 645
606 TWV-IQKRIEQFNAFEGPSLIPQELINVFDERELELLIGGISEIDMEDWKXKDYRS 664
646 CTPDSNIVKFWKAVEPDEDERARLLOFTVGSRRVLPQGFALQGAAGPRLFTIHIDA 705
665 YSENDQIIKWFELMDENSKSRLLQFTTGTSRIPVNGFKDLQSGDGRPKFTIERAGE 724
706 CTNNLPKARTCFNRIDIPYBSYEKLYEKLITATEETCGPAVE 748
725 -PNKLPKARTCFNRDLDPYTSKKDLHKLISIAVEETIGRQOE 766

RESULT 4

US-09-392-163A-4
; Sequence 4, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-392-163A-4

Query Match 38.7%; Score 1564; DB 4; Length 766;

1 MSNPGRRNGPKRLTVLCAKLVKDFRLLPFPFAKVVWDGSGQCHSTDTVKNITLDPK 60
1 MSNSAQSR----RIRVTVAADGLYKRDVFRFPFPFAVLTVDGE-QTHTTTAIKTLNPFY 55
61 WNOHYDLYIGKSDSVTISVWNNHKHKKGAGFLGCVRLLSNAINRLKDTGYQL--DLC 118
56 WNTFEVNDNSTIAIQVFDOKKF-KKKGQGLGVINLRVGDVLDLAIGDDEMLTRDLK 114
119 KLGPNDNDTVRGQIVVSLQ----- 137
115 K--SNENTVVGKIIINLSTTAQSTIQVPSSAAGARTQRTSITNDPOSSKSSVSRRNP 172
138 -----SDRIGTGQVDCSERLDN-----DLPDGEERRFASGRIOYLNHITRTQ 184
173 SSRAGSPTDAPAAASPASSEPRFTSSFDQYGLRFPGWERRTDNLGRTYVVDNTRSTT 232
185 WERTP-----RPASSYSGFRPLSCFVDENTPISGTNGATCGQSSDPLAERVRQRH 238
233 WIRPNLSSVAGAAALHLSAS--SANVTEGVQPSSNAA-----RTEASVLV---- 279
239 RNTMSTHLTPDPEGYEQRTQOGQYVFLHTGTGVSTWHDPRVPRDLNIN----- 292
280 ----SNATTAGSGELPPGWEQYTPGWRPYVDHNTRITTTWDPRRQYIRSYGPPNAT 335
293 -----CEELGELPPGWEIENATGVVYVDHNTRITQTDRLSANLHLVLRNQQLKDQ 347
336 IQQPVSLQGLPLPSGWEMELTNARVYVDHNTKTITWDDPRLPSSL-----DQ 384

Query Match 38.7%; Score 1564; DB 3; Length 766;
Best Local Similarity 41.1%; Pred. No. 7.9e-138; Indels 132; Gaps 18;
Matches 338; Conservative 121; Mismatches 232;

1 MSNPGRRNGPKRLTVLCAKLVKDFRLLPFPFAKVVWDGSGQCHSTDTVKNITLDPK 60
1 MSNSAQSR----RIRVTVAADGLYKRDVFRFPFPFAVLTVDGE-QTHTTTAIKTLNPFY 55
61 WNOHYDLYIGKSDSVTISVWNNHKHKKGAGFLGCVRLLSNAINRLKDTGYQL--DLC 118
56 WNTFEVNDNSTIAIQVFDOKKF-KKKGQGLGVINLRVGDVLDLAIGDDEMLTRDLK 114
119 KLGPNDNDTVRGQIVVSLQ----- 137
115 K--SNENTVVGKIIINLSTTAQSTIQVPSSAAGARTQRTSITNDPOSSKSSVSRRNP 172
138 -----SDRIGTGQVDCSERLDN-----DLPDGEERRFASGRIOYLNHITRTQ 184
173 SSRAGSPTDAPAAASPASSEPRFTSSFDQYGLRFPGWERRTDNLGRTYVVDNTRSTT 232
185 WERTP-----RPASSYSGFRPLSCFVDENTPISGTNGATCGQSSDPLAERVRQRH 238
233 WIRPNLSSVAGAAALHLSAS--SANVTEGVQPSSNAA-----RTEASVLV---- 279
239 RNTMSTHLTPDPEGYEQRTQOGQYVFLHTGTGVSTWHDPRVPRDLNIN----- 292
280 ----SNATTAGSGELPPGWEQYTPGWRPYVDHNTRITTTWDPRRQYIRSYGPPNAT 335
293 -----CEELGELPPGWEIENATGVVYVDHNTRITQTDRLSANLHLVLRNQQLKDQ 347
336 IQQPVSLQGLPLPSGWEMELTNARVYVDHNTKTITWDDPRLPSSL-----DQ 384

US-09-392-205A-4
; Sequence 4, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-392-205A-4

380 RQELSQOPOGHCHRIEVSREIEFESYOVOMKORPKDLWK-RLMIKFRGEGLDYGGVA 438
557 RRKLKQNDIPNFKEMKREATVLEDSYRIMGVKEADFLKARLWTEFDGKGLDYGVA 616
439 REMWLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRMGMAVHG 497
617 REMWFLSKEMFPPYGLFYSATDNYTLQINPNSGLCHEDHLSYFKFICRVAGNAVTHG 676
498 HYIDGGTLPFYKOLLGKSTLDDMELVDPDLHNSLVWILENDITGVLDHTFCVEHNAYG 557
677 KLLDGFIRFYYKMLHKPITLHDMESVDSEYNSLRWILENDPT-ELDLRIIDEELFG 735
558 BIIQHELKPKNGKSPVNEENKKEVRLVYNNWRFLRGIEAQLALOKGNEVIPPQHLKTF 617
736 QTHQHELKNGGSEIVTNKKKYYILVQWRFVNRIOQMAFKEGFFELIPQDLIKIF 795
618 DEKELELIICGLGKIDVNDKVNRLKH-CTPDSNIVKFWKAVEFFDEEREARLLQFVT 676
796 DENELELMCGLDGVDVNDWRETKYKNGYSANHQVQWFKAVLMDSEKIRLLQFVT 855
677 GSSRVPLQGFKALOGAGPRLFTIHQIDACTNNLPKHAHTCFNRIDIPPPYESYEKLYEKL 736
856 GTSRVPNGFAELYGSGPOSFTVEQW-GTPEKLPRAHTCFNRDLDPFYESPEELWDLKQ 914
737 TAIBETCGF 745
915 MAIENTQGF 923

RESULT 6

US-09-392-205A-6

Sequence 6, Application US/08539205A

Patent No. 6001619

GENERAL INFORMATION:

APPLICANT: Beach, David H.

APPLICANT: Caligiuri, Maureen

APPLICANT: Nefsky, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/539,205A

FILING DATE: 04-OCT-1995

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: CSV-005.01

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 834 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-392-205A-6

Query Match

Best Local Similarity 33.5%; Score 1353; DB 3; Length 834;

Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;

QY 159 LPDGWEERTASRIQYLNHITRTTQWERPTRPASEYSSPG-----RPLS 203
DB 246 LPSGWEERKADKGRTYVYNNHRTTITWRPIMQLAEDGASGATNSNNHLEPQIRPRS 305
QY 204 C----FVDENTPISGNGATCGGSSDPRLAERVRSQRYNYSRTHLHTPPDLPEGYEQR 260
DB 306 LGSPTVTLAPLEGAKDSVRRRAVXDTLSNPQSPQSPYNSPKQHKVQTQSFLLPPGEMR 365
QY 261 TTQOGGVYFLHTQTGVSTWHDPRVP-----RDLNSINCEELGPPPGWEIRNTATGRVYF 315
DB 366 IAPNGRPFIDENTTITTWEDPRLAKFPVHMSKTSLSNENDLGPLPPGWEERIHLDDGRFF 425
QY 316 VDHNNRTTQFTDPRLSANLHLVLNQNLKQQQQQVSVLCPDDTECLTVPR--YKRDLV 373
DB 426 IDHNSKITQWEDPRL-----QN-----PAITGPAVYSEFEK 457
QY 374 QKLKILRQELSQQQOAGHCRIEVSREIEFESYOVOMKORPKDLWK-RLMIKFRGEG 432
DB 458 QKYDYFRKLLKXEPADIPNFKEMKLRNNIPESYRIRMSVXRPDVLKARLWIEFESEKGL 517
QY 433 DYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRMG 491
DB 518 DYGGVAREWFFLLSKEMFPPYGLFYSATDNYTLQINPNSGLCHEDHLSYFTTIGRVAG 577
QY 492 MAVFHGHYIDGGFTLPFYKOLLGKSTLDDMELVDPDLHNSLVWILENDITGVLDHTFCV 551
DB 578 LAVFHGKLLDGGFIRFYYKMLHKPITLHDMESVDSEYNSLRWILENDPT-ELDLNFCI 636
QY 552 EHNAYGBEIIQHELKPKNGKSPVNEENKKEVRLVYNNWRFLRGIEAQLALOKGNEVIPPQ 611
DB 637 DEENFGQTYQVDLKPNGSEIMVTNENKREYIDLVIQWRFVNRVQKQNAFLGFTTELLPI 696
QY 612 HLLKTFDEKELELIICGLGKIDVNDKVNRLKH-CTPDSNIVKFWKAVEFFDEERE 669
DB 697 DLIKIFDENELELMCGLDGVDVNDWROHRSYKNGYC-PNHPVIONFWKAVLLMDAEKRI 755
QY 670 RLLOFTVSSRVPLQGFKALOGAGPRLFTIHQIDACTNNLPKHAHTCFNRIDIPPPYESYE 729
DB 756 RLLOFTVTSRVPMNGFAELYGSGNPQLFTIEQWS-PEKLPRAHTCFNRDLDPFYTETFE 814
QY 730 KLYEKLTAIBETCGF 745
DB 815 DLREKLLMAVENAQGF 830

RESULT 7

US-09-392-163A-6

Sequence 6, Application US/09392163A

Patent No. 6503742

GENERAL INFORMATION:

APPLICANT: Beach, David H.

APPLICANT: Caligiuri, Maureen

APPLICANT: Nefsky, Bradley

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/392,163A

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/539,205

```

; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/070,060
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHM.70312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302.886.4889
; TELEFAX: 302.886.8221
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 854 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-09-070-060-4

Query Match          33.1%; Score 1336; DB 2; Length 854;
Best Local Similarity 36.6%; Pred. No. 2.5e+116;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24

QY      13 KURLTVLCAK-NLVKKDFRLPPFAKVVDGGGCGCHSDTWTNTLDPKNOHQHYDLVIGK 71
Db       :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
           9 QLIQTIVSAKLKNKQWFG-PSPYVEVTVD--GOSKTEKCNCNTPSKWKQLTVITVP 65
QY      72 SDSVTIISVNHNKHCHK--QGAGFLGVCLLSNAIRLKDTGVQRLLDCKLGPNDDTVR 129
Db       ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
           66 TSKLCFRWSHQTILKSDVLGLTAGLDIYEtlkNNMKLEEV---VMTLOVGDKSPTETM 122
QY     130 GOVV-----SLQSRIITGGQVV-----DCSRlFDNDLPD-----GWEE 165
Db       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
           123 GDLSVCYLDGLQVAEAVVINGETSCSBSTQNDDCGRTRDTRVSTNGSEDPVAASGENK 182
QY     166 RTASGRtQYLNHITRTOWERPTRBASEVSSPGRPLSCFVDENT-----PIs 213
Db       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
           183 RANGNSFSLSNGFKPSRPDRPRPPP---PTRRPASVNGSFSTNSDSGSTSGSLPPT 240
QY    214 GTN-----CATCG-----QSSDPRLAERRVRSRHRNYM-----242
Db       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
           241 NTNVTSTSEGATSGLIIPLTISGGSGPRELNVTSQAPLP PGWEQRV-DQHGRVYVDHV 299
QY    243 -SRHLHTPPDLPGRYQRTQQGOVYFLHTGTGVSTWDHPVP-----285
Db       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
           300 EKRTTWDRPEPLPGWERRVDDNRGRIYYVDHFTRTTTWRPTLESVRNYEQWLQRSOLQ 359
QY    286 -----RDJ-----SHINCEELGPLPGWEIEINTATGRVYFDHNNRRTQFTD 327
Db       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
           360 GAMQCFNCRFTFYGNQDLFATSQNKEDPDPLGPFPGEHEKTDSNGRVYFNHNTRITQMED 419
QY    328 PRLSANHLNLNRQNLKDQQOQQCVSLCP---DDTFCLTV---PR-----y 368
Db       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
           420 PRSOQC----LNEXFLPEGWEEMFTVDGIPIYFDHNNRATTVIDPRTGKSALDNQPfAY 475

```

SULT 8
-09-070-060-4
Sequence 4, Application US/09070060
Patent No. 5976849
GENERAL INFORMATION:
APPLICANT: Hustad, Carolyn M.
APPLICANT: Ghildyal, Namit
TITLE OF INVENTION: Human E3 Ubiquitin Protein

369 KRLVQKILR---QELSQOQOAGCHRIEVSREIEEYSRQVMKMRPKDLWRLMIK 425
476 VRDFKAKVOYFRFWCOQLAMPQ---HIKITVTRKTLFEDSFQOIMFSPODLRRLRLWI 531
426 FRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDASAVNPEHLSYFHF 485
532 FRGEGLDYGGVAREWFFLLSHEVLNPMYCLFEYAGKNYCLQINPASYINPDHLKTYFRF 591
486 VGRINGMAVPHGHYIDGGTILPFYKQLLGKSIITLDDMELVDPDLHNSLVWILENDITGV- 544
592 IGRFIAMALFHGKFIDTGFSLPFYKRIILNPKVGLKDLSDIDPEFYNSLIWKENNIEECG 651
545 LDHTFCVEHNAVGEIIQHELKPNKGSIPVNEENKKEYVRLVYVNRFLRGIBAOFLALQKG 604
652 LEMFVSVDKEILGEIKSHDLKXNGNLTVEENKEEYIRMAEWRLSRGVEEQOAFPEG 711
605 FNEVTPQHLKTLTDFEKELELIICGLGKIDVNDWKVNTLKHCTPDSNIVKWFKAVEFFD 664
712 FNEILPQOYLOVFDKALEVLLCGMQEIDLNDQORHAIYRHYTRTSKQIMWFQVKEID 771
665 EBRARLLOFVTGSSRVLPQGFALQGAAGPRLETHOIDACTNNLPKAKTCNRIDIPP 724
772 NEKMRLLQFVTGTCRLPVGGFADLMGNSGPKFCIEKVGK-ENWLPRSHTCFNRDLDP 830
725 YESYEKLYEKLJTAIETTCGFAVE 748
831 YKSYEQLEKLLFAIETEGFQGE 854

SULT 9
-09-357-746-4
Sequence 4, Application US/09357746
Patent No. 6087122
GENERAL INFORMATION:
APPLICANT: ZENECA Limited
TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
FILE REFERENCE: PHM.70312.NI
CURRENT APPLICATION NUMBER: US/09/357,746
CURRENT FILING DATE: 1999-07-21
EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
EARLIER FILING DATE: 1998-02-05
EARLIER APPLICATION NUMBER: US No. 608712209/070,060
EARLIER FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 854
TYPE: PR1
ORGANISM: Mus musculus
-09-357-746-4

Query Match
Best Local Similarity 36.6%; Pred. No. 2.5e-116;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;
13 KRLTVLCAK-NLVKKDFRLDPPEAKVVDGSGGCHSTDFVKTLDPKKNHVDLYIGK 71
9 QLGITVISAKLENKKNWFG-RSPVEIVTD--GSKKTEKCNNTNSPKMKQPTVIVTP 65
72 SDSVITISVNNHKKHKK--QGAGFLGCVRLLSNAINRLKDTGYQRLDLCKLGPNDNDTVR 129
66 TSKLCERVNSHQTLKSDVLLGTAGLDIVETLKNNNKLEEV---VMTQLVQGDKEPTETM 122
130 GQIVV--SIQSDRIGTGGVV-----DCSLFDNDLPD-----GWEE 165
123 GDSVCLDGLQVEAEVVTNGTSCSESTQNDGCRTRDDTVSTNGSEDPVAASGENK 182
166 RPTASGRIOYLNHITRTQWERPTFPASEYSPGRPLSCFVDENT-----PIS 213
183 RANGNNSPSLSNGGFKPSRPPRPPPP--PTFREPAVNGSPSTNSDGSSTGSLPPT 240
214 GTN-----GATCG-----QSSDPLAERVRVSQRHNYM----- 242

Db 241 NTNVTSTSEAGTGLIPLTISGSGRPLNTVSQAPLPGEQRY-DQGRGVYVDHV 299
QY 243 -SRTELHTPPDLPEGYEORTTQOOVYFLHTQTGVSTWHDPRVP----- 285
Db 300 EKRTTWRPEPLPGWERRVDNNGRIYYVDHFRTTTTQORPTLESVRNYEQWLQSRQLQ 359
QY 286 -----RDL-----SNINCBEGLPLPGWEIRNTATGRVYFVDHNNRTTQFTD 327
Db 360 GAMOQFNQRFYGNQDLFATSNKKEFDPLGFLPGWEKRTDSNGRYVYFVNHNTRITQWED 419
QY 328 PRLSANLHLNRRQNQLKDOOQOQWLSLCP---DDTECLTV---PR-----Y 368
Db 420 PRSQOQ-----LNEKPLPEGWEMRTVDIGYFVDHNRATYIDPRTGKSALENGQIAY 475
QY 369 KRLVQKILR---QELSQOQOAGCHRIEVSREIEEYSRQVMKMRPKDLWRLMIK 425
Db 476 VRDFKAKVOYFRFWCOQLAMPQ---HIKITVTRKTLFEDSFQOIMFSPODLRRLRLWI 531
QY 426 FRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDASAVNPEHLSYFHF 485
Db 532 FRGEGLDYGGVAREWFFLLSHEVLNPMYCLFEYAGKNYCLQINPASYINPDHLKTYFRF 591
QY 486 VGRINGMAVPHGHYIDGGTILPFYKQLLGKSIITLDDMELVDPDLHNSLVWILENDITGV- 544
Db 592 IGRFIAMALFHGKFIDTGFSLPFYKRIILNPKVGLKDLSDIDPEFYNSLIWKENNIEECG 651
QY 545 LDHTFCVEHNAVGEIIQHELKPNKGSIPVNEENKKEYVRLVYVNRFLRGIBAOFLALQKG 604
Db 652 LEMFVSVDKEILGEIKSHDLKXNGNLTVEENKEEYIRMAEWRLSRGVEEQOAFPEG 711
QY 605 FNEVTPQHLKTLTDFEKELELIICGLGKIDVNDWKVNTLKHCTPDSNIVKWFKAVEFFD 664
Db 712 FNEILPQOYLOVFDKALEVLLCGMQEIDLNDQORHAIYRHYTRTSKQIMWFQVKEID 771
QY 665 EBRARLLOFVTGSSRVLPQGFALQGAAGPRLETHOIDACTNNLPKAKTCNRIDIPP 724
Db 772 NEKMRLLQFVTGTCRLPVGGFADLMGNSGPKFCIEKVGK-ENWLPRSHTCFNRDLDP 830
QY 725 YESYEKLYEKLJTAIETTCGFAVE 748
Db 831 YKSYEQLEKLLFAIETEGFQGE 854

RESULT 10
US-09-070-060-3
Sequence 3, Application US/09070060
Patent No. 5976849
GENERAL INFORMATION:
APPLICANT: Husted, Carolyn M.
APPLICANT: Grindyal, Namit
TITLE OF INVENTION: Human E3 Ubiquitin Protein
TITLE OF INVENTION: Ligase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZENECA Pharmaceuticals, Inc.
STREET: 1800 Concord Pike
CITY: Wilmington
STATE: DE
COUNTRY: USA
ZIP: 19850-5437
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,060
FILING DATE: 30-APR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/073,839
FILING DATE: 05-FEB-1998

458 PRTGKS---ALDNGPQI-----AYVRDFKAKVQFRFWCQOLA 492
385 QQPOAGCHREIVSREEIPEESYRQWVKRPPDLMKRLMIKFRGEGLDYGVAREWLYL 444
493 MPQ---HIKITVTRKTLFEDSPQIMSFSPQDLRRRLWVIFPGEGLDYGVAREWFL 548
445 LSHEMLNPPYGLFYQSRDDIYTLQINPDGAVNPEHLSYFHFVGRIMGMVAFHGHVIGGF 504
549 LSHEVLNPMYCLFEYACKONYCLOINPASYINPDHLKYFRFRIGRIAMALFHGKIDTGF 608
505 TLPFYKOLKSKSTLDDMELVDPDLNSLVILENDITGV-LDHTFCVHNAYGEIIIOHE 563
609 SLFPYKRLNKPVLGLDLSIDPEFYNSLIWVKENNIECDLEMYFSVDEKILGKISHD 668
564 LKNGKSIPIVNEENKKEYRVLRYVNMVFLGIAQFLALQKGENEVIPOHLLKTDKELE 623
669 LKNGGNILVTEENKEYIRWVAEWLRSVGEQTAFFEGNEILPOOYLOVFAKELE 728
624 LIICGLKIDVNDKVNTRLKHCCTPDSNIVKMFKAVEFDEEREARLLQFVTGSSRVPL 683
729 VLLCGMQEIDLNDQWRHATYRHVARTSKQIMVFWQFVKEIDNEKEMRLIQFVTGCRLPV 788
684 QGFKALQGAAGPRLFTHQIDACTNNLPKATCFNRIDIPPEYSEKLYEKLTAIBETC 743
789 GGFADLMSNGPQKFCIEKVGK-ENWLPSSHCTCFNRDLDPYKSYEQLEKELFAIBETE 847
744 GFAVE 748
848 GFGQE 852

SULT 12

-08-630-916A-48

Sequence 48, Application US/08630916A

Patent No. 6011137

GENERAL INFORMATION:

APPLICANT: Pirozzi, Gregorio

APPLICANT: Kay, Brian K.

APPLICANT: Fowkes, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

NUMBER OF SEQUENCES: 124

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,916A

FILING DATE: 03-APR-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-203

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 896-8864/9741

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 906 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

-08-630-916A-48

Query Match 32.2%; Score 1298.5; DB 3; Length 906;

Best Local Similarity 42.0%; Pred. No. 9e-113;

Matches 271; Conservative 101; Mismatches 216; Indels 57; Gaps 9;

QY 122 RNDNDTVRGQIVVLSQSRDRITGTGGVQVDCSRLFDNDLPDQWEERRTASGRIOYLNHTR 181
DB 301 PNPNTTSLPAPATPAEGEESTSGTQQLPAAAQADALPAGWEQRELPNGRVVYVDHNTK 360
QY 182 TTQWERPTRPASEYSSGPRPLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN- 240
DB 361 TTTWERPLPGWEKTDPRGRFYVDHNT-----RITWQRPATAEYVRYVYEQWQSRNQ 414
QY 241 -----YMSRTHLTP-DLPPEGYQORTQQQVYFLHTQTGVSTWHDPV 284
DB 415 LOGAMQHSQRFYQFWSASTDHPGLPFGWEKR-QDNGRVVYVNHNTRTTQWEDPT 473
QY 285 PRDLNSINCEELGPPPGWEIRNTATGRVYFVDHNNRTTQDPRLSANLHLVLRNQNL 344
DB 474 QGMIOE-----PALPPQWEMKYTSEGVRVFDHNTTTFKDRPGFE----- 516
QY 345 KDQOQQQVVSCLPDDTECLTVPRYKRLDVQKLILRQELSQQQPAQCHCRIEVSREIFE 404
DB 517 -----SGTKQSGPAYDRSPRWKYHQFR-FLCHSNALPSHVKISVSQRTLFE 562
QY 405 ESYRQVQVVRPKDLMKRLMIKFRGEGLDYGVAREWLYLISHMLNPPYGLFYQSRDDI 464
DB 563 DSFQIMMVKPYDLRRRLYIIMRGEGLDYGVAREWFFLLSHEVLNPMYCLFEYAGKN 622
QY 465 YTLQINPDSAVNPEHLSYFHFVGRIMGMVAFHGHVIGDGTLPFYKOLKSKITLDDMEL 524
DB 623 YCLQINPASSINPDHLTVFRFGRIFAMALYHGFIDTGTFTLPFYKEMLNKRTPLKDL 682
QY 525 VDPDLHNSLVMLENDITGV-LDHTFCVHNAYGEIIQHELKPKNGKSIPIVNEENKKEYR 583
DB 683 IDPEFYNSIVMIKENNLEECGLELYFQDMELGKVTTHLKEGGSIRVTEENKEEYIM 742
QY 584 LYVNRFLRGIEAQLALQKGFNEVIPHLLKTFDEKELELIICGLKIDVNDKVNTRL 643
DB 743 LLTDWRFTRGVEEQTKAFLDGFEVAPLEWLRVDFEKELEMLCGMQEIDMSDWKSTIY 802
QY 644 KHCPTDSNIVWFVWKAVEFDEEREARLLQFVTGSSRVPLQGFALQGAAGPRLFTHQI 703
DB 803 RHYTKNSKQIQMFQVVKEMDNKRIELLOFVTGTCRLPVGFGFAELIGSNGPQKFCIDKV 852
QY 704 DACTNNLPKATCFNRIDIPPEYSEKLYEKLTAIBETCQFAVE 748
DB 853 GKET-WLPSSHCTCFNRDLDPYKSYEQLEKELFAIBETEFGQE 906

RESULT 13

US-08-630-916A-46

Sequence 46, Application US/08630916A

Patent No. 6011137

GENERAL INFORMATION:

APPLICANT: Pirozzi, Gregorio

APPLICANT: Kay, Brian K.

APPLICANT: Fowkes, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

NUMBER OF SEQUENCES: 124

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley, Hoag & Eliot
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,904B
FILING DATE: 23-MAY-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 874 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
247-904B-8

Query Match	13.2%	Score 534.5	DB 2	Length 874
Best Local Similarity	25.1%	Pred. No. 4.9e-41		
Matches 185	Conservative 117	Mismatches 275	Indels 161	Gaps 23
QY	105	NRLKDTGQRLDLCKLGN---	DNDTVRGQIVVSLQSDRIGTG-----	GOVVDCS 152
Db	202	SRIGDSOGDNNLQKLGPDVSDIADIR-RVTRLLSNEKIETAFNALVYLSPNVECD	260	
QY	153	RLFNDNLPDGWEERTASGRIOYLNHITRTQWERPTRPASEYSSGPRPLSCFVDENTPI	212	
Db	261	LTYHN-----	VYSRDPNLYNLF--LIIGNENPLHSPEYLEWALPLPKCKAMSKLPL	308
QY	213	SGTNG-----	ATCQSSDPRLAERVRVSQRHRYMGRTHL-----	247
Db	309	AAQGKLRLMSKYNADQIRMMETFOQLITYKVISNEFNSRLNLFNFSRLNVDDDAIV	368	
QY	248	-----	HTPPD-----LPE-----	CYEORTTQCGOYVELH 271
Db	369	AASKCLMYYANVVGGEVDTNNEBDDPEEPSESELTLQELLGEERNKKGVVDPLE	428	
QY	272	TQTGVSTWHPRVPRDLSNINCDELGPLPGWGEIRNTATGRV-----	YFVDHNRT 322	
Db	429	TELGVKT-----	LDCKK--PLIPFEFNEPLNEVLMDKDYTFPKVETENKF	474
QY	323	TQTFDPRL--SANKLHLVNLQNOLKQOQQOQVSVLCDDTECLTVPRYKRDVLQKLIKLR	380	
Db	475	SFMTCPFLNAVTKNLGLYDNRIMYSERRITVL-----	-----	509
QY	381	QELSQOQPOAGHCRIEVSRREEIFEBS--YRQWKMVRPKDLKRLMIKFRGEGGLDYGV	437	
Db	510	YSLVQGQQLNPFYLRKVRRHIIIDDLVALVLEMANENPADLKQYIVSFEFGQGVDEGV	569	
QY	438	AREWLYLLSHEMLNPFYGLFQXSRDDIYTLQINPDSAVNPEHLSYFHVPGVIMGMAVPHG	497	
Db	570	SKEFPQLVVEEINPDIGNFTYD-ESTKLFWNPSSP--ETEGQFTLIGIVLGLAYNN	625	
QY	498	HYIGGTTLPFYKQLGKSLITLDDMELVDDPDHLNSLWILE--NDITGVLDHTFCV-EHN	554	
Db	626	CILDVHPFMVYRKLAKGKGLFVLDGDSHPVLYQSLKOLLEYGVNVEDMMITFQISQTN	685	

SULT 14
-08-247-904B-8
Sequence 8, Application US/08247904B
Patent No. 5981699
GENERAL INFORMATION:
APPLICANT: Rolfe, Mark
APPLICANT: Eckstein, Jens W.
APPLICANT: Draetta, Giulio

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

% protein - protein search, using sw model

on on: February 20, 2004, 15:27:02 ; Search time 33.0523 Seconds
(without alignments)
4738.482 Million cell updates/sec

itle: US-10-009-945-4

arfect score: 4038

sequence: 1 MSNPGRRNGPVKRLTVLC.....EKLYEKLLTAETTCGFAVE 748

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

sarched: 801455 seqs, 209382283 residues

otal number of hits satisfying chosen parameters: 801455

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : Published Applications AA:**

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:**
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:**
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:**
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:**
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:**
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:**
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:**
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:**
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:**
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:**
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:**
- 12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:**
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:**
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:**
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:**
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:**
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:**
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	4031	99.8	748	12	US-10-021-660-81
2	3939.5	97.6	735	12	US-10-313-955-2
3	3001	74.3	722	15	US-10-097-534-14
4	2775	68.7	514	12	US-09-764-875-819
5	1564	38.7	766	12	US-10-313-955-4
6	1530.5	37.9	832	12	US-10-032-585-7296
7	1453.5	36.0	869	15	US-10-128-714-8162
8	1370.5	33.9	911	15	US-10-205-823-279
9	1354.5	33.6	725	15	US-10-185-050-136
10	1354.5	33.5	854	15	US-10-205-823-277
11	1354.5	33.5	927	15	US-10-097-534-15
12	1353.5	33.5	995	15	US-10-097-534-9
13	1353.5	33.5	995	15	US-10-205-823-275
14	1353	33.5	834	12	US-10-313-955-6
15	1340.5	33.2	759	15	US-10-128-714-3162

16	1338.5	33.1	854	12	US-10-287-218-3	Sequence 3, Appli
17	1301	32.2	739	12	US-10-374-979-89	Sequence 89, Appli
18	1301	32.2	739	15	US-10-097-534-10	Sequence 10, Appli
19	1301	32.2	739	15	US-09-919-039-235	Sequence 235, App
20	1298.5	32.2	870	15	US-10-097-534-12	Sequence 12, Appli
21	1298.5	32.2	906	15	US-10-185-050-48	Sequence 48, Appli
22	1209	29.9	474	11	US-09-774-639-371	Sequence 371, App
23	1209	29.9	474	11	US-09-969-730-249	Sequence 249, App
24	1134.5	28.1	683	15	US-10-185-050-46	Sequence 46, Appli
25	1134.5	28.1	684	15	US-10-097-534-11	Sequence 11, Appli
26	1050	26.0	1562	15	US-10-097-534-13	Sequence 13, Appli
27	1019.5	25.2	380	15	US-10-307-956-32	Sequence 32, Appli
28	1008.5	25.0	375	15	US-10-307-956-31	Sequence 31, Appli
29	869	21.5	733	15	US-10-097-534-16	Sequence 16, Appli
30	835.5	20.7	1094	15	US-10-043-487-300	Sequence 300, App
31	835.5	20.7	1488	15	US-10-043-487-285	Sequence 285, App
32	775.5	19.2	277	10	US-09-925-300-1628	Sequence 1628, Ap
33	733	18.2	358	15	US-10-268-036-5	Sequence 5, Appli
34	682	16.9	276	10	US-09-925-300-1527	Sequence 1527, Ap
35	665	16.5	125	9	US-09-764-870-384	Sequence 384, App
36	665	16.5	125	12	US-09-764-875-1110	Sequence 1110, Ap
37	665	16.5	125	15	US-10-125-540-384	Sequence 384, App
38	499	12.4	823	15	US-10-097-534-25	Sequence 25, Appli
39	445.5	11.0	1050	15	US-10-097-534-24	Sequence 24, Appli
40	445.5	11.0	1050	15	US-10-097-534-28	Sequence 28, Appli
41	445.5	11.0	1054	15	US-10-097-534-29	Sequence 29, Appli
42	422	10.5	1024	15	US-10-097-534-30	Sequence 30, Appli
43	407.5	10.1	4861	10	US-09-919-497-70	Sequence 70, Appli
44	407.5	10.1	4861	15	US-10-097-534-26	Sequence 26, Appli
45	407.5	10.1	4861	15	US-10-146-473-49	Sequence 49, Appli

ALIGNMENTS

RESULT 1

US-10-021-660-81
; Sequence 81, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis.
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 81
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-81

Query Match 99.8%; Score 4031; DB 12; Length 748;

Best Local Similarity 99.9%; Pred No. 0;

Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSNPGRRNGPVKRLTVLCANLVKDDFFRLPPFAKVVVDGSGQCHSTDTVXNTLDPK 60

Db 1 MSNPGRRNGPVKRLTVLCANLVKDDFFRLPPFAKVVVDGSGQCHSTDTVXNTLDPK 60

Qy 61 WNHQYDLVIGKSDSVTISVWNHKKIHKQAGAGCGVLLSNAINRLKDTGYQRDLCKL 120

Db 61 WNHQYDLVIGKSDSVTISVWNHKKIHKQAGAGCGVLLSNAINRLKDTGYQRDLCKL 120

121 GPNNDIVRGQIVVLSQSRDRIGTGQVVDGSRFPDNDLPDQWEERRTASGRIOYLNHIT 180
121 GPNNDIVRGQIVVLSQSRDRIGTGQVVDGSRFPDNDLPDQWEERRTASGRIOYLNHIT 180
181 RTQWERPTPASEYSPGRLPSCFVDENPTISGTNGATCGQSSDPRLAERRVRSQRHN 240
181 RTQWERPTPASEYSPGRLPSCFVDENPTISGTNGATCGQSSDPRLAERRVRSQRHN 240
241 YMSRTHLHTPPDLPEGVEQRTTQGGVFLHTQGVSTWHDPRVDRDLSNINCELGELP 300
241 YMSRTHLHTPPDLPEGVEQRTTQGGVFLHTQGVSTWHDPRVDRDLSNINCELGELP 300
301 PGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQLDQOQQQVSLCPDDT 360
301 PGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQLDQOQQQVSLCPDDT 360
361 ECLTVPRYKRDVLQKILROELSQQOQAGHCRIFVSREEIFEESYRQVMKRPDLWK 420
361 ECLTVPRYKRDVLQKILROELSQQOQAGHCRIFVSREEIFEESYRQVMKRPDLWK 420
421 RLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDRDIYTLQINPDSAVNPEHL 480
421 RLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDRDIYTLQINPDSAVNPEHL 480
481 SYFHFVGRINGMAVFGHYIDGGFTLPFYKQLGKSIITLDDMELVDPDLHNSLVILEND 540
481 SYFHFVGRINGMAVFGHYIDGGFTLPFYKQLGKSIITLDDMELVDPDLHNSLVILEND 540
541 ITGVLDTHTFCVEHNAYGEIIOHELKPNKSGSIIPVNEENKKEYVRLVYVNRFLRGIEAQLA 600
541 ITGVLDTHTFCVEHNAYGEIIOHELKPNKSGSIIPVNEENKKEYVRLVYVNRFLRGIEAQLA 600
601 LQKGFNEVIFQHLKLTDFDEKELELIICGLGKIDVNDKVNTRLKHCCTPDSNIVKWFKA 660
601 LQKGFNEVIFQHLKLTDFDEKELELIICGLGKIDVNDKVNTRLKHCCTPDSNIVKWFKA 660
661 EFPDEERRARLLQFVGTSSRVPLOGFKALQGAAGPRLFTIHOIDACTNNLPKANTCFNRI 720
661 EFPDEERRARLLQFVGTSSRVPLOGFKALQGAAGPRLFTIHOIDACTNNLPKANTCFNRI 720
721 DIPPYSEYKLYEKLJTAIEETCGFAVE 748
721 DIPPYSEYKLYEKLJTAIEETCGFAVE 748

SUJT 2
-10-313-955-2
Sequence 2, Application US/10313955
Publication No. US20030199036A1
GENERAL INFORMATION:
APPLICANT: Beach, David H.
Nefsky, Bradley
Caligiuri, Maureen
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY, ROAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/313,955
FILING DATE: 05-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/392,163

FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/539,205
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-313-955-2

Query Match 97.6%; Score 3939.5; DB 12; Length 735;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;

QY 1 MSNPGRRNGPVKRLTLVLCANLVKDDFFALPDPFPAKVVDGSGGCHSTDTYKNTLDPK 60
Db 1 MSNPGRRNGPVKRLT-----GLPDPFPAKVVDGSGGCHSTDTYKNTLDPK 47
QY 61 WNOHYDLVIGKSDSVTISVNNKHKKQAGFLGCVRLLSNAINRLKDTGYORLDLCKL 120
Db 48 WNOHYDLVIGKSDSVTISVNNKHKKQAGFLGCVRLLSNAINRLKDTGYORLDLCKL 107
QY 121 GENDNDTVRGQIVVLSQSRDRIGTGQVVDGSRFPDNDLPDQWEERRTASGRIOYLNHIT 180
Db 108 GENDNDTVRGQIVVLSQSRDRIGTGQVVDGSRFPDNDLPDQWEERRTASGRIOYLNHIT 167
QY 181 RTQWERPTPASEYSPGRLPSCFVDENPTISGTNGATCGQSSDPRLAERRVRSQRHN 240
Db 168 RTQWERPTPASEYSPGRLPSCFVDENPTISGTNGATCGQSSDPRLAERRVRSQRHN 227
QY 241 YMSRTHLHTPPDLPEGVEQRTTQGGVFLHTQGVSTWHDPRVDRDLSNINCELGELP 300
Db 228 YMSRTHLHTPPDLPEGVEQRTTQGGVFLHTQGVSTWHDPRVDRDLSNINCELGELP 287
QY 301 PGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQLDQOQQQVSLCPDDT 360
Db 289 PGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQLDQOQQQVSLCPDDT 347
QY 361 ECLTVPRYKRDVLQKILROELSQQOQAGHCRIFVSREEIFEESYRQVMKRPDLWK 420
Db 348 ECLTVPRYKRDVLQKILROELSQQOQAGHCRIFVSREEIFEESYRQVMKRPDLWK 407
QY 421 RLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDRDIYTLQINPDSAVNPEHL 480
Db 408 RLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDRDIYTLQINPDSAVNPEHL 467
QY 481 SYFHFVGRINGMAVFGHYIDGGFTLPFYKQLGKSIITLDDMELVDPDLHNSLVILEND 540
Db 468 SYFHFVGRINGMAVFGHYIDGGFTLPFYKQLGKSIITLDDMELVDPDLHNSLVILEND 527
QY 541 ITGVLDTHTFCVEHNAYGEIIOHELKPNKSGSIIPVNEENKKEYVRLVYVNRFLRGIEAQLA 600
Db 528 ITGVLDTHTFCVEHNAYGEIIOHELKPNKSGSIIPVNEENKKEYVRLVYVNRFLRGIEAQLA 587
QY 601 LQKGFNEVIFQHLKLTDFDEKELELIICGLGKIDVNDKVNTRLKHCCTPDSNIVKWFKA 660
Db 588 LQKGFNEVIFQHLKLTDFDEKELELIICGLGKIDVNDKVNTRLKHCCTPDSNIVKWFKA 647
QY 661 EFPDEERRARLLQFVGTSSRVPLOGFKALQGAAGPRLFTIHOIDACTNNLPKANTCFNRI 720
Db 648 EFPDEERRARLLQFVGTSSRVPLOGFKALQGAAGPRLFTIHOIDACTNNLPKANTCFNRI 707
QY 721 DIPPYSEYKLYEKLJTAIEETCGFAVE 748
|||||

RESULT 5

3-10-313-955-4
Sequence 4, Application US/10313955
Publication No. US20030199036A1
GENERAL INFORMATION:
APPLICANT: Beach, David H.
Nefsky, Bradley
Calligiuri, Maureen

TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/313,955

FILING DATE: 05-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/392,163

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/539,205

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: CSV-005.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 766 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

-10-313-955-4

Query Match

Best Local Similarity 41.1%; Score 1564; DB 12; Length 766;

Matches 336; Conservative 121; Mismatches 232; Indels 132; Gaps 18;

1 MSNPGRRNGPVKRLTVLCAKLVKKDFRLLDPFAKVVVDGSGQCHSTTVKNTLDPK 60

1 MSNSAQGR-----KIRVTIWAADGLYKRDVFRFPDPAVLTVDGE-QTHHTTAIKKLNPY 55

61 WNOHYDLIGKSPSVTISVWNHKKHKQAGFLGCVRLLSNAINRLKDTGVQRL--DLG 118

56 WNETFEVNTDNSTIAIQVFDQKF-KKKGQGLGVINLVGCDVLDLAIGDEMTRDLK 114

119 KLOPNDNTVRGQIVVLSQ----- 137

115 K--SNENTVVGKIIINLSTTAOSTLOVPSSAASGARTQTSITNDPOSSSSVSRRNP 172

138 -----SRDRIGTGQVVDGCSRLFN-----DLPDGEERTASGRIQVNLHITRTQ 184

173 SSRAGSTRDNAPASAPSEPRFTSSFDQGRLLPPGWERNTDLNGLTYVDHNTASTT 232

185 WERTP-----RPASYSFGPRPLSCFVDENTPISGTNGATCGOSSDPLAERRVRSQRH 238

233 WIRFNLSVAGAAAALHSSAS--SANVTEGVQVQSSNAA-----RRTEASVLT--- 279

239 RNTYMRTHLTPDLPGEYEQRTQQGVVFLHTQTGVSTWHDPRVPRDLNIN----- 292

Db 280 ----SNATTAGSGELPPGWEQRYTFEGRPYFVDNTRTTTWDPRROQYRSYGGPNAT 335

QY 293 -----CEELGPLPPGWEIRNTATGRVYFVDHNNRTTQTPRSLSANHLVLRNQKQDQ 347

Db 336 1QQQPVSQLPLPGSGWEMELTNTARVYFVDHNTKTTTWDPRLPSSL-----DQ 384

QY 348 QQQQVVSCLPDDTECLTVPRYKRDIVQKLKILRQELSQ--QPOAGHCRLEVSREEIFE 405

Db 385 -----NVPOYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHFED 425

QY 406 SYRQVMKRPKDLWKRLMKFRGEEGLDYGAVREWLYLLSHENLNFYGLFQYSRDDIY 465

Db 426 SYAETMRQSATDLKRLMKFDGEDGLDYGSLREYFELLSHEMFNFYCLFEYSSVDNY 485

QY 466 TLQINPDSAVNEHLSYFHFVGRIMGMAVFGHVIDGGFTLPYPYKQLLGKSIITLDDNELV 525

Db 486 TLQINPHSGINPEHLNYFKFGRVIGLAIFHRFRVDAFFVVSFYKMLQKKVTLQDQESM 545

QY 526 DPDLNSLVWILENDITGVLDHTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLY 585

Db 546 DAERYSLVWLNDITGVLDLTSVEDNCFGEVVTIDLKPNGRNIETENKKEYVDLV 605

QY 586 VNWRFIRGIEAQLALQGFNEVIPHLLKTFDKELELICGLGKIDVNDWKVNTLKH 645

Db 606 TYW-IQKRIEQFNAPHEGSELIPQELINVDRELELLIGGSEIDMEDWKKHKDYRS 664

QY 646 CTPDGNIVKWFMAVEFDEERRARLLQFVTGSSRVLPQGFALQGAAGPRFLTTHOIDA 705

Db 665 YSENDQIKWFELMDENSKSRLQFTTGTSGRIIPWNGFKDLQSGDGRPKFTIEKAGE 724

QY 706 CTNNLPKHAHTCFNRIDIPPYSYEKLYEKLTAIETCGFAVE 748

Db 725 -FNKLPKARTCFNRLDLPFTSKKLDHKLKLSIAVEETIGFQGE 766

RESULT 6

US-10-032-585-7296
; Sequence 7296, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7296
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7296

Query Match 37.9%; Score 1530.5; DB 12; Length 832;
Best Local Similarity 38.4%; Pred. No. 1.4e-132;
Matches 338; Conservative 120; Mismatches 228; Indels 195; Gaps 21;

QY 6 RRRNGFVKRLTVLCAKLVKKDFRLLDPFAKVVVDGSGQCHSTDTVKNTLDPKWNQHY 65

Db 9 RSNNTTTINVKWAABESLYKRDVFRQDPDPFAVLTVDGS-QTKTTITAKTLAPYWNFT 67

QY 66 DLYIGKSPSV-TISVWNHKKHKQAGFLGCVR-----LLSNAINRLKDTGVQRLDLC 119

Db 68 N-FOAKEDSILVIQVFDQKF-KKKGQGLGVINLVGCDVLDLSNSEEITR--DLKX 123

QY 120 LQPNNDNTVRGQIVVLSQSRDRIGTGQGVVD----- 150

Db 124 --SNENLAVSGKIIWI-SHNENSGGVTATTTRGTGASSNNIATITSGVNNLRIGSA 180

151 -----CSRLPND--- 157
181 TTTANSTAASSDATVANGSGPTSSLPPIGQGHPTAATPGGAAGAAAGROQSSPED 240
158 ---DLPGWEERRTASGRIQYLNHIITTTOWERPTPASE----- 194
241 QYGLPLPGWEERRTDFORTYVYVHNSTTTWQRPALHOSETERGQORQSETEAERRQHRG 300
195 YSSPGR---PLSCFVDENTPISGTNGATCQSS-----DRLAERRVRORHRN 240
301 RTLPGEQSVSL-----PTGSGNSITSGNTVTVASGANTFVNPAAVASGATTS 352
241 YMSRTHLTPDLPPEGVEQRTTQOQVYFLHTGTGVSTWHDPRVPRDLNIN----- 292
353 GLG-----ELPSSGEORFTTEGREFYFVDHNTTRTTWDPRAQOYIRTPGNTTIOQ 404
293 -CEBLGPLPPEWEIRNTATGRVTVFVDHNRHTTQTPRLSANLHLVLRNQKLDQOQOQ 351
405 PVSQGLPLPSGWEMLNTARVTVFVDHNTKTTTWDDPRLPSSL-----DQ----- 449
352 VVSLCPDTECLTVPRYKRDVLQKILROELSOQOQO-----AGHCRIEVSREEEFESY 407
450 -----NVPOYKDFPRKRVYFR-----SQPALRILPGQCHIKVREDHIFEDSY 492
408 ROVMKRPKDLWKRLMTKPEEGDLYGGVAREWLYLLSHEMLNPPYGLFQVSRDDIYTL 467
493 QEINRQTPEDLKKRLMTKFDGEBGLDYGVSREFFELSHDMFNPFYCLFEYSSHDNTL 552
468 QINPDSAVNPEHLSYHFVGRIMGMAVFHGYIDGGFTLPFYKQLLGKSLITLDDMELVDP 527
553 QINPNSGINPEHLYNFKFIVGVVGLGVFHRRLDFAFVGALYKWLKXVVLQDNEGDA 612
528 DLHNSLWILENDITGVLDTHFCVHNAGELIOHELKPKNGKSIPIVNEENKXYVRLVYN 587
613 EFRSLKWLILNDITGLDITFSAEESFEIYEVDLKPGGRDIEVTENKHEYVELITE 672
588 WRFLRGIEAQLALQKGFNEVIVPOHLKTFDEKELELIIICGLKIDVNDWKVTELKXCT 647
673 WRISKRVQEKAFIDGFEELIQLVNVFDERELELLIGLAEIDCEDWKGDYRGYQ 732
648 PDSNIVKFWKAVEFDEERRARLLFOVTGSRVPLQGFALQGAAGPRLFTIHOIDACT 707
733 ENDQVQWFKCINWDSEQKARLLOFTGTSRIPVNGFKDQSGDPRFTIEKAGE-A 791
708 NNLKPAHTCFNRDIDPPYSYKLYEKLATAEEICGRAVE 748
792 NQLPKSHTCNFRVLDLPYTDYESLKKOKLTLAVEETVGFQGE 832

SULT 7

-10-128-714-8162

Sequence 8162, Application US/10128714

Publication No. US20030119013A1

GENERAL INFORMATION:

APPLICANT: Jiang, Bo

APPLICANT: Hu, Wengqi

APPLICANT: Tishkoff, Daniel

APPLICANT: Zamudio, Carlos

APPLICANT: Eroshkin, Alexey M

APPLICANT: Lemieux, Sebastien M

TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

TITLE OF INVENTION: Methods of Use

FILE REFERENCE: 10182-018-999

CURRENT APPLICATION NUMBER: US/10/128,714

CURRENT FILING DATE: 2002-04-23

PRIOR APPLICATION NUMBER: US 60/285,697

PRIOR FILING DATE: 2001-04-23

PRIOR APPLICATION NUMBER: US 60/287,066

PRIOR FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: US 60/295,890

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: US 60/303,899

PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 8603

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 8162

; LENGTH: 869

; TYPE: PRT

; ORGANISM: Aspergillus fumigatus

US-10-128-714-8162

Query Match 36.0%; Score 1453.5; DB 15; Length 869;
Best Local Similarity 36.2%; Pred. No. 2,1e-125;
Matches 332; Conservative 114; Mismatches 229; Indels 243; Gaps 21;

QY 11 PVKRL-TVLCAKLVKQDFR-----LPDPPAKVVVDGSGQCH 48
Db 15 PASSRLCRVIAADGKYKRDVFRKSYLAAILVVRILTSGAFAGPDPFPFATVAGGE-QTH 73
QY 49 STDVTNLTDPKXNQHLYIGKSDSVTISVMNHKKIHKQAGFLGCVRL----- 100
Db 74 TTSVIKTLNPNYWNEMFWRVNEDESILAIQIDQKF-KKKQOQFLGVINVRIGDVIDLQ 132
QY 101 -----SNAINRLK-----DTGQRLDLCGLPNDNDTVRGQIVVSL----- 136
Db 133 MGGDGBSLPIRHSKDVSRILTFRLLLLETEMLTRDLKK--SNDNLVWHGKLIINLSTNLST 190
QY 137 -----QSRDRIGTGGGVDCS----- 152
Db 191 PNTNQANGLHRSHVSSSTSSGLVPQVAPSSSHPAASGTAPVDPASNPSPNQPVPSTR 250
QY 153 -----RLFDND-----LPDGEERTASGRIQYLNHITTTQ 184
Db 251 PASTAAPASAGAAVSNHSGRTNLSSPESQGRLPAGWERREDNLGRTYVVDHNTTTT 310
QY 185 WERTPAPASEYSPGRPLSCFVDENTPISGTNGATCGOSSPRLAERRVRSORHNTM-- 242
Db 311 W--TRPSSNYNEHAQ-----RSQREANMOLERRAHSRLP 344
QY 243 -----SRTHLHTPP-----DLPEGYEORTTQOQVY 268
Db 345 EDTGTANSPLPESQQAHTPPAGGSANAVSMATGATTACTAGELPGWEQRTTPEGPY 404
QY 269 FLHTGTGVSTWHDPRVPRDL-----SNIN-----CEBLGPLPGWEIRNTATGRVY 314
Db 405 FVDHNTTRTTTWDPERQOYIRMYGQNGANGTNTTQQOQVPSQLGPLSGWEMRLNTARVY 464
QY 315 FVDHNTTRTTQTPRLSANLHLVLRNQKLDQOQOQVVSILCPDDTECLTVPRYKRLVQ 374
Db 465 FVDHNTKTTWDDPLPSSL-----DQ-----GVQYKDRFR 497
QY 375 KUKILRQELSOQOP-----OAGHCRIEVSREEIPESYRQVQMKRPKDLWKRLMKFRGEE 430
Db 498 KLIYFR-----SQPALRIMSQQCHVKVRNNIFEDSYAETMRQASDLKKRLMKPDGED 552
QY 431 GLDYGVAEWLYLLSHEMLNPPYGLFOYERDDIYTLQINPDSAVNPEHLSYHFVGRIM 490
Db 553 GLDYGGLSGERFFLLSHEMFNPFYCLFEYSADHNTYTLQINPHSGVNPHEHLNPKFGRVY 612
QY 491 GNAVPHGYIDGGFTLPFYKQLGKSLTDDMELVDPDLHNSLWILENDITGVLDHTFC 550
Db 613 GLAIFHRRLDLSFFIGAFYKWLKVKVSLQDMEGVDEDLHRLNTWTWNDIEGVLELTF 672
QY 551 VEHNAYGELIOHELKPKNGKSIPIVNEENKXYVRLVNVWFLRGIEAQLALQKGFNEVIP 610
Db 673 VDDEKFGERRTIDLPKGGDRIPVTNENKAEYVRLVTEWKIVKRVVEQFNAMFGFNEPIP 732
QY 611 QHLKTFDEKELELIIICGLKIDVNDWKVNTLKHGCTPDSNIVKWFKAVEFDEERRAR 670
Db 733 ADLVNVFDERELELLIGIADIIVDDMKKHTDVRGYQESDEVIONFWKIVRSWDAQKSR 792
QY 671 LLOFTVGTSSRPVLOGFKALQGAAGPRLFTIHOIDACTNNLPKHAHTCFNRDIDPPYSYK 730
Db 793 LLOFTVGTSSRPVNGFKDQSGDGRPRFTIEK-SGDPALPKSHTCNFRVLDLPYKSYET 851

731 LYELKLTAEIBETCGFAVE 748
: : : : :
852 LEHKWSIAVEITLGFQGE 869

RESULT 8

1-10-205-823-279
Sequence 279, Application US/10205823
Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert

APPLICANT: Monahan, John E.

APPLICANT: Endegé, Wilson O.

APPLICANT: Gannavarapu, Manjula

APPLICANT: Gorbatcheva, Bella

APPLICANT: Hoersch, Sebastian

APPLICANT: Kamatkar, Shubhangi

APPLICANT: Monsey, Angela M.

APPLICANT: Glatt, Karen

APPLICANT: Zhao, Xumei

APPLICANT: Anderson, Dustin

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND

METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

THERAPY OF PROSTATE CANCER

FILE REFERENCE: MRI-044

CURRENT APPLICATION NUMBER: US/10/205,823

PRIOR FILING DATE: 2002-07-25

PRIOR FILING DATE: 2001-07-25

PRIOR FILING DATE: 2001-07-25

PRIOR FILING DATE: 2001-08-22

PRIOR FILING DATE: 2001-08-22

PRIOR FILING DATE: 2001-09-25

PRIOR FILING DATE: 2001-09-25

PRIOR FILING DATE: 2001-12-12

PRIOR FILING DATE: 2002-03-05

NUMBER OF SEQ ID NOS: 455

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 279

LENGTH: 911

TYPE: PRT

ORGANISM: Homo sapiens

1-10-205-823-279

Query Match

Best Local Similarity 33.9%; Score 1370.5; DB 15; Length 911;

Matches 324; Conservative 119; Mismatches 234; Indels 207; Gaps 28;

24 LVKXDFRPL---DPAK---YVDGQCH---STDVTKNTLDPKWNQHYDLYIGKSD-S 74

69 LPTSDELFLPGICDPYVGLSLYVADENRELALVQTKTKTLNPKWNEFFYRVNPSNR 128

75 VTSVWNHKKHKKQAGFLGCVR-----LLSNAINRLKUTGY 112

129 LFEVFDENLRTDD---FLQGVDVFLSHLPTDPTMRPYTFKDFLLRPSHRSRVKGF 185

113 QRLDLCKLGNNDTVRGQVTSLSQSDRIGTGGVDCSRLFDND-----L 159

186 LRLKWAYMPXNG---GQDEENSQDQDDMEHGEVVD-----SNDASQHELPPL 235

160 PDGWEERTASGRIQVNLHTRTQWERT-----RPA 192

236 PGWESEKVDNLGRTYVNNHNTTQWHRFSLMDVSESDDNIRQINQEAHRRFRSRHI 295

193 SEYSSP-----GRFLSCFDVNTPISGTN-----GATCQSSDPR-LAE---RR 232

296 SEDLEPESGDPVPEWET-ISEEVNIAGDSLGALPPFPASPSRTPSELSELSRR 354

233 VR-----SQHRNVMRTHL-----HTPPD----- 252

355 LQITPDNNGEQFSLIQREPFSSRLRSCSVTDAVBQGLHLPFGAKDSPVRAVKDTLGNPQ 414

QY 253 -----LPEGVEQRTQGGQVYFLHTOTGVSTWHDPRVP-----RD 287
Db 415 SPQSPYNSPKQHVKVTSQFLPPGWEHIAFNAPRPFDFIDHNTKTTTWEDPLKFPVHMRS 474
QY 288 LSNINCEELGPLPPGWEIRNTATGRVYFVDHNNRTQFTDRLSANLHLVLNRQNLKQD 347
Db 475 KTSLNPDILGPLPPGWEERIHLDRGTFFYIDHNSKITQWEDPRL-----QN----- 519
QY 348 QQQQVSLCPDDTECLTVPR--YKRDVLQKILRQELSQQQPQAGHCRIEVSREEEFEE 405
Db 520 -----PAITGPAVPSREFKQDYFRKCLKKPADIPNRFEMKLRHNNIFEE 566
QY 406 SYRQVMKMRPKDLWK-RLMIXFRGBEGLDYGVAREWLYLLSHMLNPPYGLFOYSRDDI 464
Db 567 SYXRMSVXRPDLVKARLWIEFESEKGLDYGVAREWFFLLSKEMFNPYGLFRYSATDN 626
QY 465 YTLQINPDASAV-NPEHLSYFHFVGRIMQMAVFGHYIDGGFTLFPYKQLGKSTLDDME 523
Db 627 YTLQINPNSGLCNEDHLSYFTFGRVAGLAVFHGKLLDGGFFIRFPYKMLGKQITLNDME 686
QY 524 LYDPELHNSLVWILENDITGVLDHTFCVHNAYGBELIQHELKPNKSIIPVNEENKKEVVR 583
Db 687 SVDSYNSLKWILENDT-BLDMFCIDEENFGCTTQVDDLKPNSEIMVTNENKREYID 745
QY 584 LYVNRFLRGIEBAQFLALQKGFNEVIPOHLLKTFDEKELELIIQGLKIDVNDWKVNTRL 643
Db 746 LVIQWRFPVNRVQKMAFLGFTTELLPIDLIKIFDENELELLMCGLDGVDVNDWRQHSY 805
QY 644 K-HCTPDPSNVKWFKAVEPFDERRARLLQFVTSRVPLOGFKALQGAAPRLRLETH 701
Db 806 KNGYC-PNHPVIQWFKAVLLMDAEKIRLLQFVTGSRVPMNGFAELYSGNSGQPLFTIE 864
QY 702 QIDACTNNLPKATCFNRIDIPPEYSEKLYEKLTLTAIETCGF 745
Db 865 QWGS-PEKLPRAHTCFNRLDLPYETPEDIREKLLMAVENAQGF 907

RESULT 9

US-10-185-050-126

Sequence 126, Application US/10185050

Publication No. US20030071577A1

GENERAL INFORMATION:

APPLICANT: Pirozzi, Gregorio

Fowkes, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

NUMBER OF SEQUENCES: 233

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/185,050

FILING DATE: 28-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/826,516

FILING DATE: 03-Apr-1997

ATTORNEY/AGENT INFORMATION:

NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 19,872

REFERENCE/DOCKET NUMBER: 1101-208-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 126:

SEQUENCE CHARACTERISTICS:
LENGTH: 725 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: unknown

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 126:

-10-185-050-126

Query Match 33.6%; Score 1355.5; DB 15; Length 725;

Best Local Similarity 43.7%; Pred. No. 28-116;

Matches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;

137 QSRDRIGTGQVDCSRIFDN---DLPDGWEERTASGRIQVLYNHITTTTOWERPTRPAS 193

112 RASSTVTGGEPTPSVAYVHTTGLPSGWEERKDAKGRYYVNNHNRITTTWTRPIMQLA 171

194 EYSSPG-----RPLSC---FVDENTPISGTNGATCGQSSDPRLAERVR 235

172 EDGASGATNNHLEPQIRPRSLSSPTVLSAPLEGAKDSPVRAVKDTLSNPSPQ 231

236 QRHNMTSRTHLTPDLPPEGVEORTQOGVYFLHTGTGVSTWHDPRVP-----RDLSN 290

232 PSPYNSPKQHKVTSQFLPGWEMRIAPNGRPFIDHNTKTTTWEDPRLKPPVHMRKTS 291

291 INCEELGPIPPGWEIRNTATGRVYFVDHNRRTTQTDPRLSANLHLVLRNQLKQDQQQ 350

292 LNPNDLGPLPGWEERHLDGRFTYIDNSKITQWEDPRL-----QN----- 333

351 QVSLCPDDECTLVPR--YKDLVQKILRLQELSSQOQPGAGHCRIVSREEIFESYR 408

334 -----PAITGPAPVYSREFQKYDYFRKLLKPADIPNRFEMKLRHNNIFEYSYR 383

409 QVMKMRPKDLWK-RLMIKFRGEGLDYGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTL 467

384 RINSVKRPDVLKARLWIEFESKGLDYGVAWEFFLLSKEMFPYGLFYBSATDNYTL 443

468 QINPDASV-NPEHLSYFHFVGRINGMAVFGHIDGGFTLPFYKQLLGSITLDDMELVD 526

444 QINPNSGLCNEHLSYFTFGRVAGLAVFGHKLDDGGFTIRFFYKQWMLGKQITLNDMESVD 503

527 PDLHNSLVILENDITGVLDHTFCVEHNAYGEIIQHELKPKNGKIPVNEENKKEYYRLV 586

504 SEYNSLKWILENDPT-ELDLMFCIDEENFGQYQVDLKPNGSEIMVTNENKREYIDLVI 562

587 NWRFLRGIEAQFLAQGFNEVTPHLLKTDFEKELELIICGLGKIDVNDKVNTRLK-- 644

563 QWRFVNRVQKQNAFLGFTTELLPIDLKIIPDENELELLMGLGDVDVNDWRQHSIYKNG 622

645 HCTPDSNIVKFWKAVFEFFDEERRARLLQFVTGSSRVPLOQFKALOGAAGPRLFTIHOID 704

623 YC-PNHPVIOFWKAVLLMDAKRILQLQFVTGTSRVPMPNGFABLYGSGNGPQLFTIEQWG 681

705 ACTNNLPKATCFNRIDIPYSEYKLEKLLTALEETCGF 745

682 S-PEKLPKATCFNRLDLPYETTFEDLQKLLMAVENAQGF 721

SULT 10

-10-205-823-277

Sequence 277, Application US/10205823

Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert

APPLICANT: Monahan, John B.

APPLICANT: Endege, Wilson O.

APPLICANT: Gannavarapu, Manjula

APPLICANT: Gorbacheva, Bella

APPLICANT: Hoersch, Sebastian

APPLICANT: Kamatkar, Shubhangi
APPLICANT: Monsey, Angela M.
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: PastSeq for Windows Version 4.0
SEQ ID NO 277
LENGTH: 854
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-277

Query Match 33.5%; Score 1354.5; DB 15; Length 854;
Best Local Similarity 43.7%; Pred. No. 3.2e-116;
Matches 280; Conservative 98; Mismatches 100; Indels 63; Gaps 13;

QY 137 QSRDRIGTGQVDCSRIFDN---DLPDGWEERTASGRIQVLYNHITTTOWERPTRPAS 193

DB 241 RASSTVTGGEPTPSVAYVHTTGLPSGWEERKDAKGRYYVNNHNRITTTWTRPIMQLA 300

QY 194 EYSSPG-----RPLSC---FVDENTPISGTNGATCGQSSDPRLAERVR 235

DB 301 EDGASGATNNHLEPQIRPRSLSSPTVLSAPLEGAKDSPVRAVKDTLSNPSPQ 360

QY 236 QRHNMTSRTHLTPDLPPEGVEORTQOGVYFLHTGTGVSTWHDPRVP-----RDLSN 290

DB 361 PSPYNSPKQHKVTSQFLPGWEMRIAPNGRPFIDHNTKTTTWEDPRLKPPVHMRKTS 420

QY 291 INCEELGPIPPGWEIRNTATGRVYFVDHNRRTTQTDPRLSANLHLVLRNQLKQDQQQ 350

DB 421 LNPNDLGPLPGWEERHLDGRFTYIDNSKITQWEDPRL-----QN----- 462

QY 351 QVSLCPDDECTLVPR--YKDLVQKILRLQELSSQOQPGAGHCRIVSREEIFESYR 408

DB 463 -----PAITGPAPVYSREFQKYDYFRKLLKPADIPNRFEMKLRHNNIFEYSYR 512

QY 409 QVMKMRPKDLWK-RLMIKFRGEGLDYGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTL 467

DB 513 RINSVKRPDVLKARLWIEFESKGLDYGVAWEFFLLSKEMFPYGLFYBSATDNYTL 572

QY 468 QINPDASV-NPEHLSYFHFVGRINGMAVFGHIDGGFTLPFYKQLLGSITLDDMELVD 526

DB 573 QINPNSGLCNEHLSYFTFGRVAGLAVFGHKLDDGGFTIRFFYKQWMLGKQITLNDMESVD 632

QY 527 PDLHNSLVILENDITGVLDHTFCVEHNAYGEIIQHELKPKNGKIPVNEENKKEYYRLV 586

DB 633 SEYNSLKWILENDPT-ELDLMFCIDEENFGQYQVDLKPNGSEIMVTNENKREYIDLVI 691

QY 587 NWRFLRGIEAQFLAQGFNEVTPHLLKTDFEKELELIICGLGKIDVNDKVNTRLK-- 644

DB 692 QWRFVNRVQKQNAFLGFTTELLPIDLKIIPDENELELLMGLGDVDVNDWRQHSIYKNG 751

QY 645 HCTPDSNIVKFWKAVFEFFDEERRARLLQFVTGSSRVPLOQFKALOGAAGPRLFTIHOID 704

DB 752 YC-PNHPVIOFWKAVLLMDAKRILQLQFVTGTSRVPMPNGFABLYGSGNGPQLFTIEQWG 810

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/313,955
FILING DATE: 05-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/392,163
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/539,205
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
10-313-955-6

Query Match 33.5%; Score 1353; DB 12; Length 834;
Best Local Similarity 44.6%; Pred. No. 4.3e-116;
Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;
159 LPDGEERTASGRIOYLNHITRTQWERTRPASEVSSPG-----RPLS 203
246 LPSCWEERKDKAGTYVYNNHRTTWTIRIMQLAEDGAGSATNNHILSPQRRRS 305
204 C---FVDENPISGTGATCGSDPPLAERRVRSQRHRYMSRTHLTPPLPGYER 260
306 LSSPTVTLAPLEGAKDSPVRAVKOTLSNPQSPSPYNSPKQKVTQSFLPPGWEHR 365
261 TTQGVVFLHTQTGVSTWHDPRVP-----RDLNINCEBGLPLPGWEIRTAGRVYF 315
366 IAPNGRPFIDHNTKTTTWEDPRKLPVFNRSKTSLSNLDLPLPGWEIRHLGRIFY 425
316 VDHNRRTQTFDPLSANLHLVLRNQKLDQOQQVSLCPDDTDECLTVPR--YKRDIV 373
426 IDHNSKITQWEDPRL-----QN-----PAITGPAPVYSREPK 457
374 QKLILQELSOQOPQAGHCEIEVSRREIEFESYRQVMKRPKDLWK-RLMKFRGEGL 432
458 QKYDYFRKLGKPKADIPNRPEMKLHNNIEFESYRIMSVKRPDVLKARLWIEPESEGL 517
433 DYGVVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMG 491
518 DYGVVAREWFLLSKEMFNPPYGLFEYSATDNYTLQINPNSGLCNEDHLSYFTFGRVAG 577
492 MAVFHGYIDGGFTLPYKQLGKSTLDMELVDPDLNLSLWILENDITGVLDHTFCV 551
578 LAVFHGKLLDGGFFIRPYPKWLKQKITLNDMESVDSEYTNLSKLWILENDPT-ELDLMEFCI 636
552 EHNAVGTIIQELKPNKSIIPVNEENKKEVRLVYVNRFLRGIEAQFLALQGFNEVTPQ 611
637 DEENFGQYQVDLKPNGSELTWNENKREVIDLVQWRFNVRQKMAFLSGFTLELPI 696
612 HLLTFDEKLELLICGLKIDVNDKVNTRLK--HCTPDSNIVKFWKAVFFDSERRA 669
697 DLKIFDENELELLMCLGPDVDNDRQHSIYKNGYC-PNHVPVQWFWKAVLLMDAEXRI 755
670 RLLQFVTSRVPVLPQKALOGAGRPLFIHQIDACTNNLPAKTCFNRIDIPPEYSYE 729
756 RLLQFVTSRVPVNGFAEYVYSGNGFQFTIEQWGS-PEKLPRATCFNRDLPLPYETFE 814
730 KLYEKLLTAIBETCGF 745
815 DLREKLLMAVENAQGF 830

RESULT 15
US-10-128-714-3162
; Sequence 3162, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3162
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3162

Query Match 33.2%; Score 1340.5; DB 15; Length 759;
Best Local Similarity 41.9%; Pred. No. 5.3e-115;
Matches 279; Conservative 88; Mismatches 164; Indels 135; Gaps 12;
Qy 159 LPDGEERTASGRIOYLNHITRTQWERTRPASEVSSPGRLSCFVDENPISGTNGA 218
Db 153 LPAGWERREDNLGRTYVDHNTTITW---TFPSNYNEHAQ----- 191
Qy 219 TCQSSDDPLAERRVRSQRHRYM-----SRTHLTHPP----- 251
Db 192 -----RSQREANMQLERRAHOQRMLPEDRTGANSPLNPESQQAHTPPAGGSANAVSMA 246
Qy 252 -----DLPEGYEORTTQGVYFLHTQTVSTWHDPRVPRDL-----SNIN----- 292
Db 247 TGATTAGTCELPGWEQRTTPEGRYFYVDHNTTITWDPFRQYIRMYGQANGTWTTI 306
Qy 293 -----CEELGPLPGWEIRNTATGRVYFVDHNRRTTQTFDPLSANLHLVLRNQKLDQOQ 348
Db 307 QQQPVSQLGFLPSGWEMLTNTARVYFVDHNTKTTTWDDPLPSSL-----DQ- 354
Qy 349 QQQVWSLCPDDTDECLTVPRYKEDLVQKLILRQELSQQP-----QAGHCRIEVSSEIFE 404
Db 355 -----GVPOYKDFRKLIIYPR-----SQPALRIMSGCHVKVRRNNIFE 394
Qy 405 ESYRQVMKRPKDLWKRLMKFRGEGLDYGVAREWLYLLSHEMLNPPYGLFOYSRDDI 464
Db 395 DSYAEMRQASDLKKRLMKFDGSDGLDYGLSREFFLLSHEMFNPFYCLPYSADN 454
Qy 465 YTLQINPDSAVNPEHLSYFHFVGRINGMAVHGHYIDGGFTLPYKQLGKSTLDDMEL 524
Db 455 YTLQINPHSGVNPENHLNFKFGRVVGVAIFHRFLDLSFFFGAFYKMWLRKKVSLQDMEG 514
Qy 525 VDPDLHNSLVWILENDITGVLDHTFCVNEHNAVGTIIQELKPNKSIIPVNEENKKEVY-- 582
Db 515 VDEDLHRLNLTWLDNDIEGVLELTFESVDEKEFGERRTIDLKPGGRDIPVTNENKAEYVDC 574
Qy 583 -----RLVYVNRFLRGIEAQFLALQGFNEVTPQHLKLTFFDEKEL 622

575 YLEPLAAEDDQKINTNGFFRLVTEWKIVKRVVEQFNFMVSGFNELIPADLVNVFDEREL 634
623 ELJICG:GKIDVNDWKVNTLKHCTPDSNIVKFWKAVEFFDEERRARLLOFVTGSSRVP 682
635 ELLIGGIADIDVDWKKHIDYRGYQESDEVIQNFVKIVRSWDAEQKSLLOFTTGTSRIP 694
683 LOGFKALQGAAGPRLFTIHQIDACTNNLPAHTCFNRIDIPPYESYEXLYEKLJTAIEET 742
695 VNGFKDLOQSGDGPFRFTIEK-SGDPAALFKSHTCFNRDLDPYKSYETLEHKVSIIVEET 753
743 CGFAVE 748
754 LGRGQE 759

arch completed: February 20, 2004, 15:37:46
b time : 36.0523 secs

49744
robable ubiquitin-protein ligase [imported] - Neurospora crassa
;Alternate names: protein B24B19.160
;Species: Neurospora crassa
;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
;Accession: T49744
;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
;Reference number: Z25022
;Status: preliminary
;Molecule type: DNA
;Residues: 1-815 <SCH>
;Cross-references: EMBL:AL556192; GSPDB:GN00116; NCSP:B24B19.160
;Experimental source: BAC clone B24B19; strain OR74A
;Genetics:
;Gene: NCSP:B24B19.160
;Map position: 6
;Introns: 11/1, 24/1, 59/2, 110/1, 783/2
;Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
;239-276/Domain: WW repeat homology <WW1>
;334-371/Domain: WW repeat homology <WW2>
;393-430/Domain: WW repeat homology <WW3>
Query Match 37.7%; Score 1521; DB 2; Length 815;
Best Local Similarity 38.8%; Pred. No. 5.3e-102; Indels 170; Gaps 20;
Matches 332; Conservative 124; Mismatches 229;
16 LTVLCAKLVKDFRPPDPFAKVVVDSGQCHSTDTVKNLDPKKNQHYDLY-----I 69
9 LFVIAADGLYKRDVERPDPFAVATNGE-QTKTQVSKRTLPYWNHFDPLQIORSKV 67
70 GKSDSVTISVNNKKHKQAGLGVCLVLSNAINRLKDTGYQL-DLCKLGNNDND 126
68 NEDSILAVQVFDQKF-KKKDQGLGVINRIGSVIELAPDAEDMLTRDLK--STDNL 124
127 TVRQGVVSLQ-----SRDRI-----GTG 145
125 VVHGKLIINLTNLATMSRLGPPSPSSRSLTTPQSSVTSNDRANERSSAMSGNGTA 184
146 GOVVDCSR-----LFDND-----LPDG 162
185 NNMTLARSLASLVSSSTAPTPTNGTAPTNPSTLVPACARHHSSTLSPFDSMGRLPAG 244
163 WEERTASGRIQYLNHTRTQWERPTRPASEVSSPGRPLSCFVDENTPI----- 212
245 WEREDHLGRYYVDHNSRTSNRPT-----GTGAENRTAENTQVERQQRNRTL 297
213 -----SGNGATCGOSSDPLAERRVRSQRHNWMSRTHLTP--PDLPEGYQRITQQGQ 266
298 PEDRTGAN-----SPTLQQQQAATAAATMMHTGATTPTGELPAGWEQRTPPEGR 349
267 VYFLHTQTVSTWEDPRVPRDL-----SNIN-----CEELGLPLPPWEIRNTATGRV 313
350 PYFVDHNRITTTWDPDRQYIRMYGQNNNTGTIOQPVSQLGLPLSGWENRLNTARV 409
314 YFVDHNRITQFTDPLSLANLHLVLRNQLKQDQVQVSLCPDTECLTVPRYKRLV 373
410 YFVDHNRITTTWDPDRPLPSL-----DQ-----NVFQYKDRFR 442
374 QKLKILACLSQQQPOQAGHCRIVSREEIFESYRVQWQMRPKDLWRLIMKFRGEGLD 433
443 RLLYFRSQ-PAKRIMSGQCHIKVRSHIFEDSFABEISQASATDLKXLMIKFDGEDGLD 501
434 YGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHLSVFHFVGRIMGA 493
502 YGGLSREFFLLSHEMFNPYCLFEYSADHNTYTLQINPHSGINPEHLYFFYFGRVGLA 561
494 VFCHYIDGFTLPFFYKQLLGKSIITLDDMELVDPDLHNSLVMLENDITGVLDHTFCVEH 553
562 IFHRFIDAFIFIGALYKMWLGKAVSLADMEGVADPFRSLQWMLDNDITDVLDTFTFSD 621
554 NAYGEIITQELKNGKSIIPWNEENKKEYVRLVYVNRFLGIEAQFLAQKGFNEVIPHQL 613

Db 622 ERFGVITEEDLIPNGRIATVNNKKIYVELVMVKWAKRIKRIQQQFRAFQDGHFELIPOOL 681
QY 614 LKTDEKELELLICGLGKIDVNDKVNTRLKCHCTPDSNIVKWFKAVFPPDEERRARLIQ 673
Db 682 INVDERELLEGIAEIDVDKXKHTDYRGYTESDEVIFQFWQTVRSWDGEGQKSRLLQ 741
QY 674 FVTGSSVPLOQKAGLQAGPRLFTIHQIDACTNLPKATCFNRIDIPPYESVEKLYE 733
Db 742 FTTGTSRIPVNGFDLQSGDGRPRFTIEKAGEIT-NLPRKATCFNRLDLPFYKSLMLQQ 800
QY 734 KLLTAIEETCGFAVE 748
Db 801 KLLTAIEETMGFQE 815
RESULT 4
S43217
ubiquitin-protein ligase (EC 6.3.2.19) RSP5 - Yeast (Saccharomyces cerevisiae)
N:Alternate names: E6-AP-like protein ubiquitin ligase; protein YER125w; PUB1 protein
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S43217; S50628; S70050
R:Mulligan, J.R.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor
submitted to the EMBL Data Library, February 1993
A:Reference number: S30812
A:Accession: S43217
A:Molecule type: DNA
A:Residues: 1-809 <MUL>
A:Cross-references: GB:U18916; EMBL:U11119; NID:G1384128; PIDN:AAC03223.1; PID:G603364
R:Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of *S. cerevisiae* cosmid 9781, 8198, 9115, 9981, and 1ambc
A:Reference number: S50628
A:Accession: S50628
A:Molecule type: DNA
A:Residues: 1-809 <DIE>
A:Cross-references: EMBL:U18916; NID:G1384128; PIDN:AAC03223.1; PID:G603364; GSPDB:GNC
R:Hein, C.; Springael, J.Y.; Volland, C.; Hagenauer-Teapla, R.; Andre, B.
Mol. Microbiol. 18, 77-87, 1995
A:Title: NP11, an essential yeast gene involved in induced degradation of Gap1 and Fur
A:Reference number: S70050; MUID:96154942; PMID:8596462
A:Accession: S70050
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-101 <HEI>
C:Genetics:
A:Gene: SGD:RSP5; PUB1; NP11; MIPS:YER125w
A:Cross-references: MIPS:YER125w; SGD:S0000927
A:Map position: 5R
C:Function:
A:Description: Involved in endocytosis of Gap1 protein and Fur4 protein; binds and ub;
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C:Keywords: ligase
F:229-266/Domain: WW repeat homology <WW1>
F:331-368/Domain: WW repeat homology <WW2>
F:387-424/Domain: WW repeat homology <WW3>

Query Match 36.8%; Score 1484; DB 1; Length 809;

Best Local Similarity 37.1%; Pred. No. 2.5e-99;

Matches 324; Conservative 125; Mismatches 224; Indels 200; Gaps 20;

QY 11 PVKRLRLTVLCANLVKDPFRLPDPFAKVVVDSGQCHSTDTVKNLDPKKNQHYDL-YI 69

Db 2 PSISVSLVAESLYKRDVFRSPDPFAVLTIDGY-QYKSTSAAKKTLNPNYWNFTFKDDI 60

QY 70 GKSDSVTISVNNKKHKQAGLGVCLVLSNAINRLKDTGYQL-DLCKLGNNDND 118

Db 61 NENSILITQVFDQKF-KKKDQGLGVINRIGSVIELAPDAEDMLTRDLK 119

QY 119 KLGPNDNDTVRQIVV-----SLSQRDRIGTGGQVDCSR--- 153

Db 120 K-SNDGMVSGRLIVLKLSPSSPHSQAPSGHTASSTNTSTTTTNGHST-SSTNHS 177

Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
Reference number: Z23034
Accession: T46412
Status: Preliminary
Molecule type: mRNA
Residues: 1-820 <AAA>
Cross-references: EMBL:AL137469
Experimental source: adult testis; clone DKFZp434P2422
Genetics:
Gene: GDB:NEDD4
Note: DKFZp434P2422.1

Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
Keywords: ligase
142-179/Domain: WW repeat homology <WR1>
342-379/Domain: WW repeat homology <WR2>
393-430/Domain: WW repeat homology <WR3>
489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 33.8%; Score 1364.5; DB 2; Length 820;
Best Local Similarity 37.1%; Pred. No. 1.2e-90;
Matches 315; Conservative 115; Mismatches 221; Indels 199; Gaps 25;

50 TDVTGNTLDPKWNCHYDLYGKSD-SVTISVNNHKKIKQAGFLGCVR-----98
12 TKTIKTLNPKWNEEFVRVNPSPNHRLLFEVDENLRTDD---FLGQVDVFLSHLPTED 68
99 -----LLSNAINRLKDTGYORLDLCKLGPNDNDVTRGQVWLSQGRDRIGTG 146
69 PTMERPTFKDPLLRPSRSHKSRVGLRWKAYMPKNG-----GQDEENSQDQDMEHW 123
147 QVDCSLRFND-----LPDGHEERTASGRIQYLNHTRTOWERPT-----189
124 EVVD-----SNDASQHQBELPPPLPPGHEEKVDNLGRYYVNNHNRRTQWHRPSLMDV 178
190 -----RPAGEYSSP-----GRPLSCFVDENTPIGNTN--216
179 SSESNNIRQINQAAHRRPSRRHSIEDLEPSEGGDVPEWET-ISEEVNIAGDSLG 237
217 -----GATCQSSDPR-LAE---RRVR-----SQRHNTMSRTHL-----247
238 LALPPPASPGRSRTSPQELSELSRLQITPDNSGEGFSLIQREPPSRLRSCSVTDAV 297
248 ---HTPPD-----LPGYQRTQOQ 266
298 EQGLHPGAKDSVYRAVKDTLNPQSPQSPNSPKQHKVTQSLPFGWEMRIAPGR 357
267 VYFLHTGTGSTVTHDPRVP-----RDLNINCBELGFLPPGWEIRNTATGRVYFVDHNR 321
358 PFFIDHNTKITTWEDPLKFPVHMRKSLNPNLGLPLPGWEIRHLDGRTFYIDHNSK 417
322 TQFTDPLRLSANLHLVLRNQLKQDQOQVVSCLPDDTECLTVPR--YKRLVQKLKIL 379
418 ITQWEDPL-----QN-----PATGPAVPSRFFKQYDIF 449
380 RQELSQOQOQAGHCHREIVSREEIFEESYRQVMKMRKDLWK-RLMKRGEGLDYGVA 438
450 RKLLKPADIPNRFEMKLRHNNIFEESYRIMSKPDLVKARLAFEFSEKGLDYGVA 509
439 REWLYLLSHEMLNPNYGLFOYSRDDYITQINPDASV-NPEHLSYFHFVGRIMGAVFHG 497
510 REWFFLLSKEMFNFPYGLFYSATDNTYLTQINPNSGLCNEDHLSYFTFGRVAGLAFHG 569
498 HYIDGGTLPFYKQLGKSLTLDMLVDPDLHNSLVWILENDITGLVDHTPCVEHNAYG 557
570 KLLGGFIRFYKWLKGLQITLNDMESVSEYNSLKNILENDPT-ELDLMECIDENFG 628
558 EIIHGLKPKNGSLPNNENKBYVLYVNNRFLRGIEAQFLALQKGFNEVIFQHLKTF 617
629 QTYQVDLKPNGSEIMVTNENKREYIDIVQWRVNRVQKMAFLGFTLLPIDLIKIF 688
618 DEKELELIICGLKIDVNDKVNTRLK--HCTPDNSIVVMFWKAYEFFDEBRARLLQFV 675

Db 689 DENELELLMCGLDGVNDVNRQHSIYKNGYC-PNHVFIQWFKAVLLMDAEKRIRLLQFV 747
QY 676 TGSSRVLPQGFKALQGAAGPRLEFTIHQIDACTNNIPKAHTCFNRIDIPPEYSEYKLYEK 735
Db 748 TGTSRVPMNGFAELYGNGPQLFTIEQWGS-FEKLPRANTCFNRLLDPPYEFEDLREKL 806
QY 736 LTAIEETCGP 745
Db 807 LMAVENAQQF 816

RESULT 7
ubiquitin ligase Nedda4 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 18-Aug-2000
C:Accession: S70642
R:Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.
EMBO J. 15, 2371-2380, 1996
A:Title: WW domains of Nedda4 bind to the proline-rich PY motifs in the epithelial Na(+)
A:Reference number: S70642; MUID:96221297; PMID:8665844
A:Accession: S70642
A:Status: Preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-887 <STA>
A:Cross-references: EMBL:U50842; NID:gl293646; PIDN:AAB48949.1; PID:gl293647
C:Genetics:
A:Gene: Nedda4
C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiq
F:54-167/Domain: protein kinase C2 region homology <KC2>
F:246-283/Domain: WW repeat homology <WM1>
F:402-439/Domain: WW repeat homology <WM2>
F:459-496/Domain: WW repeat homology <WM3>
F:555-881/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 33.6%; Score 1355.5; DB 2; Length 887;
Best Local Similarity 37.1%; Pred. No. 6e-90;
Matches 314; Conservative 131; Mismatches 248; Indels 153; Gaps 25;

QY 14 LRLTVLCAKMLVKQDFRLLPDPFAKV---VDGSGQCHSTDTVNTLDPKWNQHYDLY- 68
Db 77 VRVYTAGIAGLAKDILGASDPYRVLYDPMGSLTVTQTKIKSLNPKWNEI-LFR 135
QY 69 -ICKSDSVTISVNNHKKIKQAGFLGCVR-----LLSNAIN 105
Db 136 VLPQQRHILFEVDENLRTDD---FLGQVDVLYLPTEPRMPTFKDFVLHPSH 192
QY 106 RLKDTGYORLDLCKLGPNDNDVTRGQVVSQSRDRIGTGQVVD---CSRL-----FDN 157
Db 193 KSRVKGYLKLMYLPKNGSD---ENADQAEELBPGWVLDQPDAAATHLQHPPEPS 246
QY 158 DLPDGWEERTASGRIQYLNHTRTOWERPT-----RPAS 193
Db 247 PLPPGWEERQDVLGRYYVNNHESRTQKRPSPEDDLTDDENGDIQLQAHGAFTRQIS 306
QY 194 E-----YSSGRPLSCFVDENTPIGNTGATCQSSDP-----RLAE-230
Db 307 EDVGDGPNHSPENWEIVREDENTTYSG-QAVQPPSGHPDVQVRLAEELDTLRTMYGNP 365
QY 231 ---RRVRSQRH---RNYMSRT-----HLHTPDLPEGYBORTTQOQVYFLHTQT 274
Db 366 ATSQPVTSSNHSRSGSSQTCTFEQPTLPVLLPTSSGLPPGWEKQDDGRSYVDHNS 425
QY 275 GVSTWHD-----RVP---RDLNINCBELGFLPPGWEIRNTATGRVYFVDHNRIT 323
Db 426 KTTWSKPTMQDDPRSKIPAHLRGKTPVDSNDLGLPLPGWEERTHTDGRVFFINHNKKT 485
QY 324 QFTDPLRLSANLHLVLRNQLKQDQOQVVSCLPDDTECLTVPRYKRLVQKLKILQEL 383
Db 486 QWEDPRM-----QNVAITGAEP-----YSRDYKRKYEFFRKL 519
QY 384 SOQPOAGHCHREIVSREEIFEESYRQVMKMRKDLWK-RLMKRGEGLDYGVAREW 442

520 KQTDIPNKFEMKLRANILEDYSYRINGVKRADFLKARLWIEFDGKGLDYGVAREWF 579
443 YLLSHEMLNPPYGLFOYS-RDDIYTLQINPDSAV-NPEHLSYFHFVGRMGMAVEGHVI 500
580 FLISKEMFNPPYGLFYSATEDNYTLQINPNSGLCNEDHLSYFKFGRVAGNAVHGLK 639
501 DGGTLPFYKQLLKSITLDDMELVDPDLNLSLWILENDITGVLDHTFCVBNHAYGEII 560
640 DGFIRPFYKMKLQKLTLDHMSVDSYSSLRWILENDPT-ELDLRFIIDEELFGQTH 698
561 QHELKPNKGSIPVNEENKKEYVRLVYVNRFLRGIEAQFLALOKGENEVIPOHLLKTFDE 620
699 QHELKTSSEVVTNKKKXYIYVQWRVNRVRIQKMAAFKSGFELIPQDLIKIFDEN 758
621 ELELIICGLKIDVNDKVNRLKH-CTPDSNIVKFWKAVFFDEERARLLOFVTGSS 679
759 ELELLMCGLDVNDWREHTKYKNGYSNLHQVHFWKAVLMDSEKRIQLLOFVTGTS 818
680 RVPLOQFKALOGAAGPRLFTIHOIDACTNNLPKATCFNRIDIPPPYSEKLYEKLTAI 739
819 RVPNGFAELYGSGNGPOSFIVEQW-GTPDKLPRAHTCFNELDLPYPYSEFDELWDLQWAI 877
740 BETCGF 745
878 ENTQGF 883

SULT 8
3196
DB-4 ORF - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Aug-2000
Accession: I83196
Kumar, S.; Tomooka, Y.; Noda, M.
Ochem, Biophys. Res. Commun. 185, 1155-1161, 1992
Title: Identification of a set of genes with developmentally down-regulated expression
Reference number: I60167; MUID:92328780; PMID:1378265
Accession: I83196
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: mRNA
Residues: 1-708 <RES>
Cross-references: GB:D10714; NID:9220508; PID:9220509
Genetics:
Gene: NEDD-4
Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
40-77/Domain: WW repeat homology <WW1>
196-233/Domain: WW repeat homology <WW2>
251-288/Domain: WW repeat homology <WW3>
347-682/Domain: ubiquitin-protein ligase homology <UBI>
Query Match 29.1%; Score 1177; DB 2; Length 708;
Best Local Similarity 40.4%; Pred. No. 3.7e-77;
Matches 260; Conservative 87; Mismatches 178; Indels 118; Gaps 17;
159 LPDGEERTASGRIOYLNHITTCWERPT-----RPAE 194
42 LPEGEERQVLGRITVYVNHESRRQWKRPSPDDLTDNDMDQLOAQAFTTRQISE 101
195 -----YSSPCRLSCFVDENTPISG-----TNGATCGQ--S 223
102 DVDGPDNRESPENWEIVREDENTVSGCAVQSPSGCHTDVTHLAENFTRLAVCGNPAT 161
224 SDPLAERVRSORHNNYMSRTH-----LHTPDLPEGVEQRTQOQGVYFL 270
162 SQP-----VTSNHSRGSGSLQTCIFEEQPTLPVLLPTSSGLPFGWEKQDGRGRSYV 215
271 HTGTGVSTWHDP-----RVPRDL-SNINCEELGPLPGWEIENATATGVYVDNHR 321
216 DHNSKTTTWSKTMQDDPRSKLPAHLRKTGTSNDLGLPLPGWEERTHTDGRVFFINNIK 275
322 TTQFTDPLRSLANHLVNRQNLKDCQOQVVSCLPDDTECLTPRYKRDVLQKILRQ 381
276 KTQWEDPRL-----QNVAITGP-----AVP-YSDYKRYEYFRR 309

382 ELSQOQFQGHCRIVSVREIEEYSYRQVMKQRPDLWK-RLMIKFRGEEGLDYGVARE 440
310 KKKQTDIPNKFEMKLRANILEDYSYRINGVKRADFLKARLWIEFDGKGLDYGVARE 369
441 WLYLISHMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRMGMAVEGHY 499
370 WPLISKEMFNPPYGLFYSATDNYTLQINPNSGLCNEDHLSYFKFGRVAGNAVHGLK 429
500 IDGGTLPFYKQLLKSITLDDMELVDPDLNLSLWILENDITGVLDHTFCVBNHAYGEI 559
430 LGDFFIRPFYKMKLQKLTLDHMSVDSYSSLRWILENDPT-ELDLRFIIDEELFGQTH 488
560 IQHELKPNKGSIPVNEENKKEYVRLVYVNRFLRGIEAQFLALOKGENEVIPOHLLKTFDE 619
489 HQHELKTSSEIYVVTNKKKXYIYVQWRVNRVRIQKMAAFKSGFELIPQDLIKIFDE 548
620 KELELIICGLKIDVNDKVNRLKH-CTPDSNIVKFWKAVFFDEERARLLOFVTGTS 678
549 NELELLMCGLDVNDWREHTKYKNGYSNMHQQVHFWKAVVMDSEKRIQLLOFVTGT 608
679 SRVPLQGFKALOGAAGPRLFTIHO-----IDACTNNLPK--AHTC 716
609 SRVPMNGFAELYGSGNGPOSFIVEQWGTLLSCCEHTPASIAWTC 651

RESULT 9
T37900
probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T37900
R;Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21752
A:Accession: T37900
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-671 <RIB>
A:Cross-references: EMBL:AL117390; PIDN:CAB55856.1; GSPDB:GN00066; SPDB:SPAC1805.15C
A:Experimental source: strain 972h-; cosmid c1805
C:Genetics:
A:Gene: SPDB:SPAC1805.15C
A:Map position: 1
A:Intons: 60/2, 105/1, 639/2
A:Superfamily: WW repeat homology
F;242-279/Domain: WW repeat homology <WWR>
Query Match 25.9%; Score 1046.5; DB 2; Length 671;
Best Local Similarity 33.9%; Pred. No. 1e-67;
Matches 257; Conservative 124; Mismatches 263; Indels 113; Gaps 17;
13 KRLITVLCANLVKQDFRLLPDPFAKVVVDGSGQCHSTDFVKNVLDPKWNQHYDLYIGKS 72
7 EVQLTILHVEGLWKNGLSLKPYLLISVD-DQQFIKTNVASGTLALSNGFTQKLTVPQ 65
73 DSVITSVNHHKIHKKQAGFLGCVLLSNA---INELKDTGYQRDLCKLGPNDNDTVR 129
56 SIILLQLPDEKQ-KNETSDGFLGAAVNSFLPFNNPKDDYKTRITL----RSPGSGYR 120
130 QGIVVVSQSRDRIGTGQVVD---CSRFLDNDLPQWBERRTASGRIOYLNHITRTTQW 185
121 GS-VVCLPFRSKFLPEELPADKQICTDIIDDASGCWETRIDEFGHYVL-----170
186 ERPTPASEYSSRPLSCFVDENTPISGNGATCGQSSDPLAERVRVSRQRNMTST 245
171 -----KSP-----QLSVISAISHEKLENLTPKQLE--VFSQFLFNNQSKS 209
246 HLHTPPD-----LPEGEQRTTQOQGVYFLHTQTGVSTWHDPVPRDLNINCEELGPL 299
210 SLKINLEYKVKHLEHYPLALSVRQV-----AVEKGPL 244
300 PPGWEIRNTATGRVYVDNHRNTTQFTDPR-----LSANHLVLRQNLKDCQOQVVS 351


```
b 245 PAGWEMRLSESDYVYFVDHSTKTWTSSDPDNNVAVSDSVSTDSI-----QQINDEYQK 300
Y 352 VVSLCPDDTECLTVPRYKEDLVQKLIIRQLSQOQPAQAGHCRIEVSRREIFESYRQVM 411
b 301 IA-----YWD-----RPMVANDAQ---LQKVSRAATTFEDAYDIIS 335
Y 412 KVRPKDLWRLMIKFRGEGDYGVARWELYLSSHMLNPPYGLFOYSRDDIYTLQINP 471
b 336 KLSVSDMKKLLIRFENBDGLDYGVSREFFYILSHALFNPGLYSLFFAYADNNGYQIISP 395
Y 472 DSANVPEHLSYFHFVGRIMGMAVPHGYIDGGFTLPYKQLLGKSIITLDDMELVDPDLN 531
b 396 LSSVNPDRSRFRFVGRVGMALHRRYLDQVFLPFYKILQKPLEDVKOVDEVYZE 455
Y 532 SLAVTLENDITGVLDHTFCVHNAYGEIIQHELKPKNGKSIPIVNEENKKEVRLYVNWREL 591
b 456 SLUKIKNDVDESCLNFSVEENRFGESVTVDLLPGRNTAVNNQNMWYKALTEHKLIV 515
Y 592 RGIEAQLALQKGFNEVYIPQHLTKTFEKELELIICGLKIDVNDKWNTRLXHTCPDSN 651
b 516 TSTEQFNALGGLNELIPDSVLQIFNENELDTLLNGKRDIDVDWKRFPTYRSYETD 575
Y 652 IVKFWKAVEPFDERRARLLQFVTGSSRVLQGFKALQAGAPRFTTHQIDACTNNLP 711
b 576 IVINFWELLSWSEPKAKLLOFATGTSRPLSGKDMHSGDGRFTTEKV-GHISQLP 634
Y 712 KAHTCFNRIDIPPYESYKYLEKULTAIBETCGPAVE 748
b 635 KAHTCFNRID-PPYNSKELEPQKLTIAIQTAGPTE 671

RESULT 10
37964
robable ubiquitin ligase - fission yeast (Schizosaccharomyces pombe)
;Species: Schizosaccharomyces pombe
;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
;Accession: T37964
;Devlin, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, September 1997
;Reference number: Z21757
;Accession: T37964
;Status: preliminary; translated from GS/EMBL/DDBU
;Molecule type: DNA
;Residues: 1-3227 <DEV>
;Cross-references: EMBL:Z99531; PIDN:CAB16714.1; GSPDB:GN00066; SPDB:SPAC19D5.04
;Experimental source: strain 372h-; cosmid c19D5
;Genetics:
;Gene: SPDB:SPAC19D5.04
;Map position: 1

Query Match 21.2%; Score 858; DB 2; Length 3227;
Best Local Similarity 45.4%; Pred. No. 5e-53;
Matches 169; Conservative 74; Mismatches 123; Indels 6; Gaps 3;

/ 379 LRQLSQOQPAQAGHCRIEVSRREIFESYRQV-MKVRPKDLWRLMIKFRGEGDYG 437
> 2858 LHEEAKEQYPP--LNTIVARDHVLDYSRALHFKDADEKESKLNHIDRDEGVAGV 2915
/ 438 AREWLXLLSHMLNPPYGLFOYSRDDIYTLQINPDSANVPEHLSYFHFVGRIMGMAVPHG 497
> 2916 TREWLQVLARQMFDPYALFVPGATTFHPNRDSSVNPDLHSPFKFTGRIIGKALYDG 2975
/ 498 HYIDGGFTLPYKQLLGKSIITLDDMELVDPDLHNSLWILENDITGVLDHTFCVHNAYG 557
> 2976 RLDDCHFSRAVYKHMHLRSVSVKDIESLDPYKSLVWMLNDITDIITEFAVEKDVFG 3035
/ 558 EIIQHELKPKNGKSIPIVNEENKKEVRLYVNWRELRIEAOFLALQKGFNEVYIPQHLTKTF 617
> 3036 EKTVDLIPNRIPIVTELNKQYVNMVDYKLRBSVDQQLKSLDGFSDIIPSHLIQIF 3095
618 DEKELELIICGLKIDVNDKWNTRLXHTCTPDSNIVKFWKAVEPFDERRARLLQFVTG 677
```

```
Db 3096 NEQLELLISGLPEIDIDDWKNTEVHGNNVSSPQVQWTFWRAVRSFDEBERAKLQFATG 3155
QY 678 SSRYPLOQFKALQAGAPRFTTHQIDACTNNLPKHAHTCFNRIDIPPYESYKYLEKULIT 737
Db 3156 TSKVPLANGFKLEGMGSGFORFNHKSGLSVGLNRLPQSHTCFNQLDLEPYDYQLRSLMLIT 3215
QY 738 AIBETC---GPA 746
Db 3216 AINEGSEGFGA 3227

RESULT 11
T49799
N/Alternate names: protein [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C/Accession: T49799
R/Schulte, U.; Aign, V.; Hohenseel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatur
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49799
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-3839 <SCH>
A/Cross-references: EMBL:AL356834; GSPDB:GN00116; NCSP:B11B22.10
A/Experimental source: BAC clone B11B22; strain OR74A
C/Genetics:
A/Gene: NCSP:B11B22.10
A/Map position: 6
A/Introns: 16/3; 2607/1; 2623/1; 2658/1; 2845/1; 2987/2; 3204/3; 3694/1; 3809/1

Query Match 21.0%; Score 846.5; DB 2; Length 3839;
Best Local Similarity 42.9%; Pred. No. 4.4e-52;
Matches 168; Conservative 79; Mismatches 128; Indels 17; Gaps 4;

QY 372 LVQKLKIL-----ROELSQQOQPAQH-----CRIEVSREEIFESYRQVMKVRPKDL 418
Db 3448 LVKNPKVLEFDNKNRNFNSVHSKYQOTRHSPPPLQLQVRREHVFHDSFRSLYYKKADEL 3507
QY 419 -WKRLMKFRGEGDYGVARWELYLSSHMLNPPYGLFOYSRDDIYTLQINPDSANVP 477
Db 3508 KFGKLNIRFGSEGVDAGVTREWFQVLSRQMFDPYVLFVVPVSSDRITTFHPKLSPIIND 3567
QY 478 EHLISYFHFVGRIMGMAVPHGYIDGGFTLPYKQLLGKSIITLDDMELVDPDLHNSLWIL 537
Db 3568 EHLPPFKFTGRIIGKALYEGRLLECVFSRAVYKILGKPKSVKDMESFDPDYKSLVWML 3627
QY 538 ENDITGVLDHTFCVHNAYGEIIQHELKPKNGKSIPIVNEENKKEVRLYVNWRELRIEAO 597
Db 3628 ENDITDIITEFSDVDDVFGVVDLIENGNIPIVTEENKHEYVRLIVEHKLITSVKDKQ 3687
QY 598 FLALQKGFNEVYIPQHLTKTFDEKELELIICGLKIDVNDKWNTRLXHTCTPDSNIVKMF 657
Db 3688 MKAFLTGPHIIEPELIAFNQLELELLISGLPDIDIDDKANTEYHNYSAGAPQIQFW 3747
QY 658 KAVEPDEERRARLLQFVTGSSRVLQGFKALQAGAPRFTTHQIDACTNNLPKHAHTCF 717
Db 3748 RAVRSPDKELAKLQFVTGTSKVPLNGFKELEGNVGSRENIHRDYGSKDRLPSSHTCF 3807
QY 718 NRIDIPPYESYKYLEKULITAI---EETCGFA 746
Db 3808 NQLDLPEYENYETLASQLLKAITAGSDYFGFA 3839

RESULT 12
T01491
ubiquitin-protein ligase homolog F1707.15 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 04-Mar-2000
C/Accession: T01491
R/Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; I
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel,
```

A;Residues: 1-4056 <STO>
A;Cross-references: GB:A005173; NID:g8778329; PIDN:AAF79338.1; GSPDB:GN00141
C;Genetics:
A;Gene: F14J16.10
A;Map position: 1

Query Watch 19.5%; Score 803; DB 2; Length 4056;
Best Local Similarity 40.0%; Pred. No. 6.9e-49;
Matches 177; Conservative 62; Mismatches 143; Indels 60; Gaps 7;

QY 363 LTVPRYKRDVLVKLIILRQLSQQQPQ--AGHCRIEVSREEIFEESYRVQMWRPKDLWK 420
Db 3617 LKAPRL-IDFNNKAYFRSIRHQHQBCHISGPLRISRRAVYLEDYSNQLMRSPQDLKG 3675
QY 421 ELMIKEREGELDGGVAREWLVLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Db 3676 RUNVFQREBGIDAGGUTREWYQLLSRVFDKGLLFTTVGNDA-TFQPNPNVSQTDEHL 3734
QY 481 SYTFEVGRIMGMVAFHGHIYIDGGFTLPFFYKQLIGKSITDDMELVDPLHNSLVWLLEND 540
Db 3735 SYFEKFGVMVAKLFDQGLDVVFTSFYKHILGVKVTHDIEADVDPYYKNLKWLEND 3794
QY 541 TCVULDHDTF-----CVE---HNAYGEIIQH 562
Db 3795 VSDIIDLTLFSMDADEEXHILYTEVRLAMCFCLFPFCPIPXCHGCIETILLSLMKKVTDY 3854
QY 563 ELKPNGKSI PVNEENKEVYRLVYNWFRLRGTEAQFLALQKGFNEVIPOHLLKTDFDEKL 622
Db 3855 ELKPGRNIRVTBTKHEYVDLVAGHILNATRPQINAELEGFNELIPREIVSVIFNDKEL 3914
QY 623 ELLICGLGKID-----VNDKVNTRLKHCTPDPSNVIKVFWKAVPFDEER 667
Db 3915 ELLISGLPEIDCKLDSIQYLCAYPILLDKANTETS YTAGSPVIHFWEVWKAFSKED 3974
QY 668 RARLLOFVTGSSRVPLOGFKALOGAAGPLFIHQIDACTNNLPKHAHTCFNRIDIPPYES 727
Db 3975 WARFLQFVTGTSKVPLEGFKALQGISGPORLQHKAYGAPELRPSAHTCFNQLDLPYQS 4034
QY 728 VEKLYEKLLTAIBETC---GFA 746
Db 4035 KEQLQRLLLAIHASGFGFA 4056

RESULT 14
S69625
hypothetical protein YDR457w - Yeast (*Saccharomyces cerevisiae*)
C;Species: Saccharomycetes cerevisiae
C;Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C;Accession: S69625
R;Dieckrich, F.S.
Submitted to the EMBL Data Library, August 1995
A;Description: The sequence of *S. cerevisiae* cosmid 9410, 8035, 8166, and 9787
A;Reference number: S69554
A;Accession: S69625
A;Molecule type: DNA
A;Residues: 1-3268 <DIE>
A;Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64910.1; PID:g927738; MIPS:
C;Genetics:
A;Gene: SGD:TOM1
A;Cross-references: SGD:S0002865; MIPS:YDR457w
A;Map position: 4R

Query Match 19.5%; Score 789; DB 2; Length 3268;
Best Local Similarity 45.7%; Pred. No. 5.1e-48;
Matches 164; Conservative 59; Mismatches 132; Indels 4; Gaps 2;

QY 384 SQCOFAQHCRIEVSREEIFEESYRVQ-MKMRPKOLWKRLMIKRGEGLDVGGVAREWL 442
Db 2905 NQEPK--LPITVRREQVFLSYRALFFKTNDELKNSLBITHKSGSGVDAGGVTRWEY 2961
QY 443 YLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHLSYFHFFVGRINGMAVFHGHIYDG 502
Db 2962 QVLSROMFNPDYALFLPVPSDKTTTPNRTSGINEPHLSFFKKFGIMGIKAIROCDFLDC 3021

503 GFTLFFKQLGKSIITLDDMELVDPDLHSLVWILENDITGVLDHTFCVEHNAYGEIIOH 562
3022 HFSREVYKNTLGRPVSLKOWESLDPPYKSLVWILENDITDIETFSVETDDYGEHRVI 3081
563 ELKPNKGSIIVNENKKEYVRLVNVNRFRLGIBAQFALQKGFNEVIPHOLLKTDEKEL 622
3082 NLLEGKDIIVTANKQDYKVKVEYKQLTSVKEQMDNLFVGFYALISKDLITIDEQEL 3141
623 ELICGLGKIDVNDWKVNTKRLKHTPDSNIVKFWKAVEFFDEERRARLLQFVTGSSRVP 682
3142 ELLISGLFDIVDDWKNTTYVNTATCKEVSYFMRVARSFDAERAKILLQFVTGTSKVP 3201
683 LOGFKALQGAAGRLFTTHOIDACTNNLPRKHTCFNRIDIPYVESYEKLYEKLTAIEE 741
3202 LNGFKELSGVNGVCKTSIHRDFSSERLPSSTCTFNQLNLPFYESYETLRGSLLLAINE 3260
SUSLT 15
0274
Protein: protein F45H7.6 - Caenorhabditis elegans
Species: Caenorhabditis elegans
Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
Accession: T20274; T22257
Percy, C.
Submitted to the EMBL Data Library, November 1994
Reference number: Z19245
Accession: T20274
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 1-889 <W1>
Cross-references: EMBL:Z46793; PIDN:CAA86773.1; GSPDB:GN00021; CESP:F45H7.6
Experimental source: clone C56G7
Percy, C.
Submitted to the EMBL Data Library, June 1994
Reference number: Z19538
Accession: T22257
Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: DNA
Residues: 1-889 <W12>
Cross-references: EMBL:Z34800; PIDN:CAA84325.1; GSPDB:GN00021; CESP:F45H7.6
Experimental source: clone F45H7
Genetics:
Gene: CESP:F45H7.6
Map position: 3
Introns: 30/3; 54/1; 118/1; 185/2; 213/1; 260/3; 288/2; 353/2; 454/1; 559/2; 642/2; 74
Superfamily: WW repeat homology
229-266/Domain: WW repeat homology <W1>
372-408/Domain: WW repeat homology <W2>
Query Match 17.6%; Score 709.5; DB 2; Length 889;
Best Local Similarity 27.2%; Pred. No. 4.2e-43;
Matches 194; Conservative 126; Mismatches 271; Indels 123; Gaps 19;
131 QIVVLSQSRDRIGTGGGVDCSLFDNDLPDGWEERRTAGRIQVNLNHTRTQWERPTR 190
203 EIVFSIEPKTSKSAADIMEIASSQTPESHWKTYLDAKKKFFVNHVKETRWTKEDT 262
191 PASYSFPGRLSCFVDNTPISGTNGATCGSSDPR---LAERRVRSQRHR----- 239
263 LNNNHIEPETPVHKRLSD-----RSASPRNGFITPRRTITVRSAGCPKSLDI 309
240 NYMSRTHLHTPPDLPEGYEQRTTQCGVYPLHTQTGVSTWHDFRVPRLS----- 289
310 QFORDEFKAL-----YENQDAMQ-----IYNECSVVRHAIHQKLDLPPSKFENQPL 359
290 -----NINCELGPLPGWEIRNTATGRVYFVDHNNRTTQFTDPLR-----SANLHV 337
360 FVRVNLFAITQPLPSGNEC-ITMNNRTVFLNHNKETSFYDPRIRRPETKTSRGRSV 418
338 LNRONQL-KQQQQQQVSLCPDDETC-----LTVPRYKRL 372
419 PSRSSTAHKGKIDHALISKEDLRKIAQDNFPQIAERISKLMILIERFGLAVASLANDL 478

QY 373 -----VQKL-----KILRQELSQQOPQAGHCRI--EVSREEIFEESYQVM 411
Db 479 DITLALSMLDSNTEKLAGSGDNIKMFYEDMKKEKLGKPSRLCKWKSRRLLDDAFRIIL 538
QY 412 KMRPKDLWK-RLMIKFRGBEGIDYGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTIQIN 470
Db 539 NVDPFFVLKKSRLHIRPEGELALDYGLSREFILLRSLFHPKNGYFEYBND-YHQLR 597
QY 471 PDSAVNPEHLSYFHFVGRIMGMNAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLH 530
Db 598 PRGCEETEKEKKWLLILCGRVLALAVIHRCYIDVFFFTNVFYKSLQKRPVTLMDFKESDAEFY 657
QY 531 NSLVWLTLENDITGVLDHTFCVEHNAYG-----EIQHELKPNKGSIPVNE 575
Db 658 KSNWLTLENDVVD-LEMSFVYSSVMWNGKVSIDLITSSIVSVALQABOELVPGGESQWTE 716
QY 576 ENKKEYVRLVNVNRFRLGIEAQFALQKGFNEVIPHOLLKTDFEKELELIICGLGKIDVN 635
Db 717 ANKAEPIDLMCQKKAIRGVKEPLEILLTSFNQILNDNLNLSLESSDLKRIILSGSLELDLN 776
QY 636 DWKVNTRLKHTCTPDSNI-VKWFKAVEFFDEERRARLLQFVTGSSRVPLOGFKALQGAAG 694
Db 777 DWRTNTIYKGSYSDCHIVVWFVWEVETNTQERFDLLLFTVGSSVPPFEGFSALRGNEE 836
QY 695 PRLFTTHOIDACTNNLPRKHTCFNRIDIPYVESYEKLYEKLTAIEETCGFAVE 748
Db 837 ISKFCIEKWGDAT-SFPRAHTCFNRQLQPSNTYKQQLKSKLQQAIVNGMSYSIE 889

Search completed: February 20, 2004, 15:26:57

Job time : 21.7974 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 protein - protein search, using sw model

on on: February 20, 2004, 15:25:26 ; Search time 35.0863 Seconds
(without alignments)
5501.382 Million cell updates/sec

File: US-10-009-945-4
Effect score: 4038
Sequence: 1 MSPGRRRGPVKRLTFLC.....EKLYKLLTAIEETCGFAVE 748

oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 830525 seqs, 258052604 residues

tal number of hits satisfying chosen parameters: 830525

imum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

atabase :

1: sp.archea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.ornelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.virus.*
13: sp.vertebrate.*
14: sp.unclassified.*
15: sp.virus.*
16: sp.bacteriap.*
17: sp.archesp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3041	75.3	728	11	Q8K300	Q8K300 mus musculu
2	2313	57.3	1061	5	Q9V853	Q9V853 drosophila
3	1910	47.3	376	13	Q98T84	Q98T84 xenopus lae
4	1688.5	41.3	355	11	Q8BSC0	Q8BSC0 mus musculu
5	1552	38.4	288	4	Q96DE7	Q96DE7 homo sapien
6	1418	35.1	786	3	O14326	O14326 schizosacch
7	1398	34.6	911	4	Q8NSA7	Q8NSA7 homo sapien
8	1386	34.3	955	4	Q96PU5	Q96PU5 homo sapien
9	1379	34.2	295	4	Q8NDJ8	Q8NDJ8 homo sapien
10	1378	34.1	258	11	Q9CSE3	Q9CSE3 mus musculu
11	1364.5	33.8	820	4	Q9NT88	Q9NT88 homo sapien
12	1358	33.6	887	11	Q8BGB3	Q8BGB3 mus musculu
13	1357.5	33.6	855	11	Q8CF10	Q8CF10 mus musculu
14	1355.5	33.6	887	11	Q62940	Q62940 rattus norv
15	1354.5	33.5	854	4	Q9H2M4	Q9H2M4 homo sapien
16	1354	33.5	855	11	Q9BRT9	Q9BRT9 mus musculu

17	1353.5	33.5	995	4	O43165	O43165 homo sapien
18	1353	33.5	855	4	Q8WU09	Q8WU09 homo sapien
19	1353	33.5	858	4	Q9BW58	Q9BW58 homo sapien
20	1346.5	33.3	855	11	Q99PK2	Q99PK2 mus musculu
21	1345.5	33.3	971	13	O42573	O42573 xenopus lae
22	1341.5	33.2	834	5	Q9SR64	Q9SR64 drosophila
23	1336	33.1	854	11	O54971	O54971 mus musculu
24	1333	33.0	949	5	Q9TCH4	Q9TCH4 drosophila
25	1329	32.9	862	4	Q9BY75	Q9BY75 homo sapien
26	1327	32.9	777	11	O08758	O08758 mus musculu
27	1313.5	32.5	903	4	Q96F66	Q96F66 homo sapien
28	1303.5	32.3	956	5	Q8TOR6	Q8TOR6 drosophila
29	1301	32.2	703	4	Q8451	Q8451 homo sapien
30	1301	32.2	739	4	O43584	O43584 homo sapien
31	1298.5	32.2	870	4	Q96CZ2	Q96CZ2 homo sapien
32	1298.5	32.2	870	4	O00308	O00308 homo sapien
33	1297	32.1	922	4	Q9H0M0	Q9H0M0 homo sapien
34	1297	32.1	1007	5	Q9VVI3	Q9VVI3 drosophila
35	1286.5	31.9	870	11	Q9DBH0	Q9DBH0 mus musculu
36	1281	31.7	617	11	Q8BZZ3	Q8BZZ3 mus musculu
37	1274.5	31.6	792	5	Q95XJ3	Q95XJ3 caenorhabdi
38	1271.5	31.5	794	5	Q9N2Z7	Q9N2Z7 caenorhabdi
39	1244	30.8	724	5	Q9BKW4	Q9BKW4 caenorhabdi
40	1224.5	30.3	838	5	Q9STQ0	Q9STQ0 drosophila
41	1220.5	30.2	518	5	Q8TQC8	Q8TQC8 drosophila
42	1134.5	28.1	684	4	O00307	O00307 homo sapien
43	1092	27.0	1581	4	Q9P2P5	Q9P2P5 homo sapien
44	1050	26.0	1585	4	Q9HCC7	Q9HCC7 homo sapien
45	1046.5	25.9	671	3	Q9UTG2	Q9UTG2 schizosacch

ALIGNMENTS

RESULT 1
Q8K300 PRELIMINARY; PRT; 728 AA.

AC Q8K300; 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Similar to E3 ubiquitin ligase SMURF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR ENBL; BC029097; AAH29097.1; -.
DR InterPro; IPR000008; C2.
DR InterPro; IPR001202; WW_RSP5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00832; HECT; 1.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS00237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS00020; WW_DOMAIN_2; 2.
SQ SEQUENCE 728 AA; 83083 MW; C355291B9D8AD757 CRC64;

Query Match 75.3%; Score 3041; DB 11; Length 728;
Best Local Similarity 75.0%; Pred. No. 2.5e-228;
Matches 572; Conservative 63; Mismatches 78; Indels 50; Gaps 7;

1 MSNPRRRG-PVKRLRLTVLCAKLVKIDPFRLLPDPFAKVVVDGSGQCHSTDTVKNLTDP 59
1 MSNPGTERRSSKIRLTVLCAKLVKIDPFRLLPDPFAKVVVDGSGQCHSTDTVKNLTDP 60
60 KNCQHYDLYTGKSDSVTISVNNHKKIHKQAGFLGCVALLSNAINRLKDTGYQRLDLCK 119
61 KNSQHYDLYTGKSDSVTISVNNHKKIHKQAGFLGCVALLSNAINRLKDTGYQRLDLCK 120
120 LGPNNDNTVRGQVWSLQSDRIGTGSGVQVDSRLPDNDLPDGMERRTASGRIOYLNHI 179
121 LNPSTDVAVGQVWSLQSDRIGTGSGVQVDSRLPDNDLPDGMERRTASGRIOYLNHI 163
180 TRTQWERPRPASEYSSGPRPLSCFVDENTPISGNGATCG-----QSSDPLRAE 230
164 -----EDSGPRPLSCFVEBPAYTDTGTGAAGCGNCRFVSPSDORLLV 209
231 RRVRSQSHRYM-----SRTHLTPPDLPGYEQRTTQGOVYFLHTQTGVSTWHPVPR 286
210 QRLNPEVRGPIQTQNRPHGQSPPELPGYEQRTTQGOVYFLHTQTGVSTWHPDPIPR 269
287 DLSNICEELGPIPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQNLKD 346
270 DLNSVNCDELGPPLPGWEVRSTVSGRIYFVDHNNRTTQTPDPR-----LHIMNHQCOLKE 325
347 QQQQ-QVSLCPDDTECLTVPKYKDLVOKLILKQELSQOQPOQAGHCRIVSREEIFEE 405
326 PSQPLQPLSGSVDEBELPAQRYERDLVOKLILKQELSLQOQPOQAGHCRIVSREEIFEE 385
406 SYRQVVKMPKDLKWLKMLKFRGEGLDYGVARWLYLLGHEMLNPFYGLFQVSRDIIY 465
386 SYRQIMKMPKDLKWLKMLKFRGEGLDYGVARWLYLLGHEMLNPFYGLFQVSTNIIY 445
466 TLQINPDQAVNEHLSYFFVGRINGMAVHGHIYDGGFTLPFYKQLLGKSIITLDDMELV 525
446 TLQINPDQAVNEHLSYFFVGRINGMAVHGHIYDGGFTLPFYKQLLGKSIITLDDMELV 505
526 DPLHNSLVWILENDITGVLDHFCVHNAYGEIITQHELKPKNGKSIIPVNEKKEVRLY 585
506 DPLHNSLVWILENDITGVLDHFCVHNAYGEIITQHELKPKNGKSIIPVNEKKEVRLY 565
586 VNWRFURGTEAFLAQKGFNEVIFPHLLKTFDEKLELIICGLGKIDVNDKVNTRLKH 645
566 VNWRFURGTEAFLAQKGFNEVIFPHLLKTFDEKLELIICGLGKIDVNDKVNTRLKH 625
646 CTPDSNIVKFWKAYEFDDEERARLQVFTGSSRVPLQGGKALQAGAPRLFTIHOIDA 705
626 CTVDSNIVKFWKAYEFDDEERARLQVFTGSSRVPLQGGKALQAGAPRLFTIHOIDA 685
706 CTNNLPKATCFNRIDIPPEYSEKLYEKLTAIBETCGFAVE 748
686 NTDNLPKATCFNRIDIPPEYSEKLYEKLTAIBETCGFAVE 728
QSV853 PRELIMINARY; PRT; 1061 AA.
QSV853; Q9U3W2;
01-MAY-2000 (TrEMBLrel. 13, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
CG4943 protein (E3 ubiquitin ligase) (Ubiquitin-protein ligase) (Smad-
ubiquitin E3 ligase Smuirlf).
LACK OR SMURF OR SMURF1 OR CG4943.
Drosophila melanogaster (Fruit fly).
Sukarya; Metazoa; Arthropoda; Hexapoda; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
STRAIN=Berkeley;
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

Amnatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Aahburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Ahril J.P., Achavani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu H.O.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
Evans C.A., Gocayne J.D., Amnatides P.G., Brandon R.C., Rogers Y.,
Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
Ferrieria S., Frise E., Galle R.F., Garg N.S., George R.A.,
Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
"Sequencing of Drosophila melanogaster genome."
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
Hradecky P., Huang Y., Kaminker J.S., Prochkin S.E., Smith C.D.,
Tup J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
"Annotation of Drosophila melanogaster genome."
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
Flybase;
Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

[6]
SEQUENCE FROM N.A.
Laurencin A., Hawley S.;
"Molecular cloning of a type E3 Ubiquitin ligase";
Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE FROM N.A.
Podos S.D., Hanson K.K., Wang Y.-C., Ferguson E.L.;
"The Dsmurf ubiquitin-protein ligase restricts BMP signaling spatially
and temporally during Drosophila development";
Dev. Cell 1:0-0 (2001).
[8]
SEQUENCE FROM N.A.
Liang Y.-Y., Lin X., Feng X.-H.;
"dSmurf1, a Smad-ubiquitin E3 ligase, specifically targets dpp-
activated Mad protein for degradation";
Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
-!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
EMBL; AF003802; AAF57824.3; -
EMBL; AF216521; AAF21125.1; -
EMBL; AF416571; AAL09691.1; -
EMBL; AF464851; AAM09646.1; -
HSSP; Q13526; 1PIN.
FlyBase; Fgn002906; lack.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT_domain.
InterPro; IPR002349; WW.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00397; WW; 3.
PRINTS; PR00403; IWDOMAIN.
SMART; SM00239; C2; 1.
SMART; SM00456; WW; 3.
PROSITE; PS00499; C2_DOMAIN_1; 1.
PROSITE; PS00499; C2_DOMAIN_2; 1.
PROSITE; PS00499; C2_DOMAIN_3; 1.
PROSITE; PS00499; HECT_1.
PROSITE; PS00499; HECT_2.
PROSITE; PS00499; WW_DOMAIN_1; 2.
PROSITE; PS00499; WW_DOMAIN_2; 3.
Ligase.
SEQUENCE 1061 AA; 115675 MW; 6BCC550F5129163 CRC64;
Query Match 57.3%; Score 2313; DB 5; Length 1061;
Best Local Similarity 45.1%; Pred. No. 2.5e-171;
Matches 476; Conservative 107; Mismatches 157; Indels 316; Gaps 15;
7 RRGVGVKRLVLCANLKKDFFLPDPFAKVVVDSGGCHSTDTVNTLDPKWNQHYD 66
8 RRGVGVKRLVLCANLKKDFFLPDPFAKVVVDSGGCHSTDTVNTLDPKWNQHYD 67
67 LYIGKSDSVTISVNHKKIKKQAGFLGCVRLLSNAINELKDTGQRLDCKLGNPND 126
68 LFLGIGDAITITVNNQRIKH--GSGFLGCVRIAPFNIQSLKAGQFRLDLGKLSPPDDE 125
127 TVRGVGVVLSQRDR-----IGTGGQVVDGSR--LPDNLDPGWEERRTASGRIC 175
126 LVRGQIIISLSKDGPSGNGPLAIVGSGDVRGSDSDSDSDSLPFGWEERRTDGRVY 185
176 LNHITRTQWRPRTP-----
186 VNHATKSTQWRPRTP-----
192 -----ASYSSGRLSCFV-----
206 -----
246 HSRDLSTVADSRHSTEILSVGKENTPTPTVSATTTTPGKTSNSSSSAGGRTLEQ 305
207 -----DENTPIST-----
306 RPTNEPAPTSSTTSASVRLHSNDHVKTPQHTNGHAPSTPTSGTQQNVNGNAQN 365
217 GAT-----CGSSDPLAER-----
366 GSTSGNGSGQAQPSASNGTWGTQDAATTSPSTTSPRHSQSPPTPNISPPATPSA 425

QY 232 -----RVRSQRHNTMSTHLLHTPP----- 251
Db 426 NGVHSPNANSTPAGSGGSRSYAATPQGRSSSRSSQQGGESSTRRSSRGTRNGOTS 485
QY 252 -----DLPEGYEORTTQOQGVYFLHTQTGVSTWHDP 283
Db 486 GGGGGGGGQRYASAAIAAQAARFPLDLPGYEMRTTQOQGVYFHTPGVSTWHDP 545
QY 284 VPRDL--SNINCELGPLPGWIRNATGRVYFDHNNRTTQTDPRLSANLHLVLRQ 341
Db 546 IPRDFDTQHLTDLAIGLPSGWEQRKTSGRVYFDHNNRTTQTDPRLSGSLQIMRRG 605
QY 342 N-----OLKQOQOQVVSL-----CPDDETE 361
Db 606 TVPTSAANAGTPAPSPATPSAAAAPPQATPASNATPTLTITNPPHRIVDPDLPQ 665
QY 362 CL-----TVPRYKRDVLQKILRQELSOQOQAGHCRIEVSREIFESYQVMMRPK 416
Db 666 GLLEGADLLPKYRDLVGLRALRTELQTMQPSGHCRLIEVSREIFESYRLIMKRAK 725
QY 417 DLMKRLMIKRGEGLDYGVAREWLYLASHENLPPYGLFOYSRDDIYTLQINPDSAVN 476
Db 726 DMKRLMVKRGEGLDYGVAREWLYLASHENLPPYGLFOYSRDDIYTLQINPDSAVN 785
QY 477 PEHLSYFHFVGRINGMAVFHGYIDGGFTLPYKQLGKSIITLDDMELVDPDLHNSLVI 536
Db 786 PDHLSYFHFVGRINGMAVFHGYIDGGFTLPYKQLGKSIITLDDMELVDPDLHNSLVI 845
QY 537 LENDITGLDHTFCVEHAYGEIIQHELKXGKSPVNEENKCYVRLYVNRWFLRGIEA 596
Db 846 LESNIGIIESTFVSNNSGFALVWHELKXGKSPVNEENKCYVRLYVNRWFLRGIEA 905
QY 597 QFALQKGFNEVIPHLLKTFDEKELELIICGLKIDVNDKVNTRKHKCTPDSNIVKWF 656
Db 906 QFALQKGFNEVIPHLLKTFDEKELELIICGLKIDVNDKVNTRKHKCTPDSNIVKWF 965
QY 657 WKAVEFPDEERRARLLQFVTGSSRVLPQGFKAQ---GAAGPRLFTI--QIDACTNNLPK 712
Db 966 WQVVEYSSEMRARLLQFVTGSSRVLPQGFKAQ---GAAGPRLFTI--QIDACTNNLPK 1025
QY 713 AHTCFNRIIDIPPYSEKLYEKLTAIETCGFAVE 748
Db 1026 AHTCFNRIIDIPPYSEKLYEKLTAIETCGFAVE 1061
RESULT 3
Q98TS4
ID Q98TS4 PRELIMINARY; PRT; 376 AA.
AC Q98TS4
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE E3 ubiquitin ligase Smurf2 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21107656; PubMed=11158580;
RA Zhang Y., Chang C., Gehling D.J., Hemmati-Brivanlou A., Derynck R.;
RT "Regulation of Smad degradation and activity by Smurf2, an E3
ubiquitin ligase";
RL Proc. Natl. Acad. Sci. U.S.A. 98:974-979 (2001).
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AY014181; AAG50422.1; -
DR HSSP; Q13526; 1PIN.
DR InterPro; IPR000008; C2.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00397; WW; 3.
DR SMART; SMC0239; C2; 1.

```
SMART: SM00456; WW: 3.
PROSITE: PSS0004; C2_DOMAIN_2; 1.
PROSITE: PSS0159; WW_DOMAIN_1; 1.
PROSITE: PSS0020; WW_DOMAIN_2; 3.
Ligase.
NON TER 376
SEQUENCE 376 AA; 42652 MW; CA6AB2B5F4C20F98 CRC64;

Query Match
  47.3%; Score 1910; DB 13; Length 376;
  354; Conservative 10; Mismatches 8; Indels 4; Gaps 3;

1 MSNPGRRNGPVKRLRLTVLCAKLVKKDFFR-LPDPFAKVVVDGSGQCHSTDTVKNTLDP 59
1 MSNQSRRNGPVKRLRLTVLCAKLVKKDFFGLPDPSPAKVVVDGSGQCHSTDTVKNTLDP 60
60 KWNQHYDIYIGKSDSVISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGQVRLDLCK 119
61 KWNQHYDIYIGKSDSITISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGQVRLDLCK 120
120 LQPNNDIVRQIVVLSQSRDRIGTGQGVVDCGRLFDNDLPDQWEERRTASGRIOVLNHI 179
121 LQPNNDIVRQIVVLSQSRDRIGSGGVVDCSELFDNDLPDQWEERRTASGRIOVLNHI 180
180 TRTTOWERPTPASEYSPGRPLSCFVDENTPISGTNGATCGOSSDPRLAERVRQRHR 239
181 TRTTOWERPTPASEYSPGRPLSCFVDENTPITGTNGASGQTSDPRISERRVRQRHR 240
240 NYMSTHLHTPPDLPEGVQRTTQGGVFLHTGTGVTWHDPRVPRDLNINCELGPL 299
241 NYMSTHLHTPPDLPEGVQRTTQGGVFLHTGTGVTWHDPRVPRDLNINCELGPL 300
300 PGWEIRNTATGRVYFVDHNNRTQFTDPLSANLHLVLRQN-QLKDDQQQQQVSLC-- 356
301 PGWEIRNTATGRVYFVDHNNRTQFTDPLSANLHLVLRQNQLKEQPPQVQVSLCQL 360
357 PDTECLTVPRYKDL 372
361 PDEVECLTVPRYKDL 376

MULT 4
ISCO
Q8BSC0 PRELIMINARY; PRT; 355 AA.
Q8BSC0;
01-MAR-2003 (TRENBL-rel. 23, Created)
01-MAR-2003 (TRENBL-rel. 23, Last sequence update)
01-MAR-2003 (TRENBL-rel. 23, Last annotation update)
Hypothetical HECT domain (Fragment).
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TIGSUS=Embryo;
MEDLINE=22354683; PubMed=12466851;
The FANTOM Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
Nature 420:563-573(2002).
EMBL; AK034736; BAC28813.1; -.
Hypothetical protein.
NON TER 1
SEQUENCE 355 AA; 41315 MW; 93B39B4C82F86D01 CRC64;

Query Match
  41.3%; Score 1668.5; DB 11; Length 355;
  307; Conservative 25; Mismatches 20; Indels 3; Gaps 1;

397 VSREEIFEESYQVKKRPKDLKKLWVKFRGEGLDYGGVAREWLYLLCHEMLNPPYGL 456

Db 1 VSREEIFEESYQVKKRPKDLKKLWVKFRGEGLDYGGVAREWLYLLCHEMLNPPYGL 60
Qy 457 FOYSRDDIYTLQINPDSAVNPEHLSYFPHVGRIMGMVAFHGHYIDGGFTLPFYKQLLGKS 516
Db 61 FOYSTDNITLQINPDSINPDHLSYFPHVGRIMGLAVFHGHYINGGFTVPFYKQLLGKP 120
Qy 517 ITLDDMELVDPDLHNSLWILENDITGLDHTFCVHNAYGRIIQHELKPNKGSIPVNEE 576
Db 121 IQSDLESVDPELHKSILWILENDITPVLDHTFCVHNAPGRILQHELKPNGRVPTVEE 180
Qy 577 NKKEYVRLVYNNRFLRGIEAQLALQKGFNEVIPQHLKTFDEKELELIICGLGKIDVND 636
Db 181 NKKEYVRLVYNNRFLRGIEAQLALQKGFNELIPQHLKTFDEKELELIICGLDKIDLND 240
Qy 637 WKNTNRLKCTPDSNIVKFWKAVFDEERARLLQFTGSSRVPLOGFKALQ---CAA 693
Db 241 WKNTNRLKCAVDSNIVRWFQAVETFEERARLLQFTGSTRVLPLOGFKALQGSTGAA 300
Qy 694 GPRLFTIHQIDACTNNLPKHAHTCFNRIDIPPYSEYKLYEKLTAIEETCGFAVE 748
Db 301 GPRLFTIHLIDANTDNLKHAHTCFNRIDIPPYSEYKLYEKLTAIEETCGFAVE 355

RESULT 5
Q96DE7 PRELIMINARY; PRT; 288 AA.
AC Q96DE7;
DT 01-DEC-2001 (T-EMBL-rel. 19, Created)
DT 01-DEC-2001 (T-EMBL-rel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBL-rel. 23, Last annotation update)
DE Similar to E3 ubiquitin ligase SMURF2 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009527; AA09527.1; -.
DR InterPro; IPR000569; HECT_domain.
DR Pfam; PF00632; HECT; 1.
DR SMART; SM00119; HECTc; 1.
DR PROSITE; PS00237; HECT; 1.
FT NON TER 1
SQ SEQUENCE 288 AA; 33255 MW; FE2B43E300E66537 CRC64;

Query Match
  38.4%; Score 1552; DB 4; Length 288;
  Best Local Similarity 100.0%; Pred. No. 8.4e-113; Indels 0; Gaps 0;
  Matches 288; Conservative 0; Mismatches 0;

Qy 461 RDDIYTLQINPDSAVNPEHLSYFPHVGRIMGMVAFHGHYIDGGFTLPFYKQLLGKSITLD 520
Db 1 RDDIYTLQINPDSAVNPEHLSYFPHVGRIMGMVAFHGHYIDGGFTLPFYKQLLGKSITLD 60
Qy 521 DMELVDPDLHNSLWILENDITGLDHTFCVHNAYGRIIQHELKPNKGSIPVNEEKK 580
Db 61 DMELVDPDLHNSLWILENDITGLDHTFCVHNAYGRIIQHELKPNKGSIPVNEEKK 120
Qy 581 YVRLVYNNRFLRGIEAQLALQKGFNEVIPQHLKTFDEKELELIICGLGKIDVNDWKVN 640
Db 121 YVRLVYNNRFLRGIEAQLALQKGFNEVIPQHLKTFDEKELELIICGLGKIDVNDWKVN 180
Qy 641 TRLKCHTTPDSNIVKFWKAVFDEERARLLQFTGSSRVPLOGFKALQGAAGPRLFTI 700
Db 181 TRLKCHTTPDSNIVKFWKAVFDEERARLLQFTGSSRVPLOGFKALQGAAGPRLFTI 240
Qy 701 HQIDACTNNLPKHAHTCFNRIDIPPYSEYKLYEKLTAIEETCGFAVE 748
Db 241 HQIDACTNNLPKHAHTCFNRIDIPPYSEYKLYEKLTAIEETCGFAVE 288
```



```
249 FRSRRHISELEPESEGGDVPEWET-ISEEYNIAGDSGLGLALPPPPASPSRSTRSQEL 307
228 -----LAER-----RVRS-----ORHNNYSRTHLTTPDDL 253
308 SEELSRRLQITPDNGSGQFSLQREPSRLSRSCVTDVAEQGHLPPSPVAVHTTGL 367
254 PEGYEQRITTOGOVYFHTQTGVSTWHD- -----RVPRDL 288
368 PSGEERKADAGRYYNNHNRITTTWRPIWQLAEDGASGATSNNNHLEEQIRPRSL 427
289 SNINCEBLP-----
428 SSTPTVTLAPLEGAKSPVRAVKDTLSNPQSPQSPYNSPKPHKVTSQSLFPPGWEKRI 487
308 TATGRVYFVDHNNRTTQTPRLSANHLVLRNQNLKQQQVVSCLPDD----- 359
488 APNGRPFDHNTKTTWEDFLKFFVM-----RSKTSLNPNLDGLPPLPGW 534
360 -----TECLTVPR--YKRDVQKILKILQELSQOQFP 389
535 BERTHLDRFTYIDHNSKITQWEDPRLQNEPAITGPAVPYSREPKQYDYFRKKLKPADI 594
390 AGHCRIVSREEIEESYVROWKMRPKDLWK-RLMIKFRGEGLDYGVAREWLVLSHE 448
595 PNREMKLHNNIFEESYRIMSVRPDVLKARLWIEFSEKGLDYGVAREWVFLLSKE 654
449 MLNPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMGMAVFHGHYIDGGFTLP 507
655 MFNPYGLFEYSATDNYTLQINPNSGLCNEDHLSYTFIFGRVAGLAVFHGLKLDGFFIRP 714
508 FYKOLLGKSTLDDMELVDPDLHNSLWILENDITGVLDHTCFVEHNAYGEIIOHELKPN 567
715 FYKMLGKQTLNDSVDSYNSLSKWLILENDPT-ELDLMFCDIEENFGQYQVDLKPEN 773
568 GKSIPVNEENKKEYVRLYNVNRFLRGIEAQFLALQKGFNEVPOHLLKTFDEKELELIIC 627
774 GSEIMVTNENKREVIDLVIQRFVNRVOKOMNAFLEGFTELPIDLIKFDNELELWLC 833
628 GLGKIDVNDKVNTRLK--HCTPDSNIVKFWKAVEFFDEERRARLLOFVTGSSRVPLQG 685
834 GLGVDVNDVROHSIYKNGYC-ENHPVQVFWKAVLLMDAEXRIQLQVITGSRVPMNG 892
686 FKALQGAAGPRLFTTHOIDACTNNLPKATCFNRIDIPPYESYEKLYEKLITAISETCGF 745
893 FAELYGSGNGLFTIEQWGS-PEKLPRATCTFNRDLDPPIYTFEDUREKILMAVENAQGF 951

:RESULT 9
:INDD8
:Q8NDD8: PRELIMINARY; PRT; 295 AA.
:Q8NDD8:
:01-OCT-2002 (TReMBLrel. 22, Created)
:01-OCT-2002 (TReMBLrel. 22, Last sequence update)
:01-MAR-2003 (TReMBLrel. 23, Last annotation update)
:Hypothetical protein (fragment).
:DKFZP564H223.
: Homo sapiens (Human).
:Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
:Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
:NCBI_TaxID=9606;
:[1]
:SEQUENCE FROM N.A.
:TISSE=Brain;
:Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
:Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
:EMBL: AL834242; CAB38919.1; -.
:InterPro: IPR000569; HECT_domain.
:Pfam: PF00632; HECT; 1.
:SMART: SM00119; HECTC; 1.
:PROSITE: PS50237; HECT; 1.
:Hypothetical protein.
:NON_TER 1
:SEQUENCE 295 AA; 29670 MW; BOBE04BCB1PCBC98 CRC64;
```

```
SQ SEQUENCE 295 AA; 34030 MW; 0507325127A943EA CRC64;
Query Match 34.2%; Score 1379; DB 4; Length 295;
Best Local Similarity 85.4%; Pred. No. 2,7e-99;
Matches 251; Conservative 24; Mismatches 19; Indels 0; Gaps 0;
QY 455 GLFOYSRDDIYTLQINPDSAVNPEHLSYFHFVGRIMGMAVFHGHYIDGGFTLPFYKOLLG 514
Db 2 GLFOYSTDVNLQINPDSINPDHLSYFHFVGRIMGMAVFHGHYINGGFTVPFYKOLLG 61
QY 515 KSTLDDMELVDPDLHNSLWILENDITGVLDHTCFVEHNAYGEIIOHELKPNKGSIPVN 574
Db 62 KPIQLSLESVDPELHLSLWILENDITPVLDHTCFVEHNAGRILQHELKPNGVPT 121
QY 575 EENKKEYVRLYNVNRFLRGIEAQFLALQKGFNEVPOHLLKTFDEKELELIICGLGKIDV 634
Db 122 EENKKEYVRLYNVNRFLRGIEAQFLALQKGFNEVPOHLLKTFDEKELELIICGLGKIDV 181
QY 635 NDWKVNTRLKHTCTPDSNIVKFWKAVEFFDEERRARLLOFVTGSSRVPLQGFALQGAAG 694
Db 182 NDWKSNTRLKHCVADSNIVRWQAVETFEERRARLLOFVTGSTRVPLQGFALQGAAG 241
QY 695 PRLFTTHOIDACTNNLPKATCFNRIDIPPYESYEKLYEKLITAISETCGFAVE 748
Db 242 PRLFTTHLIDANTDNLKATCFNRIDIPPYESYEKLYEKLITAISETCGFAVE 295

:RESULT 10
:Q8CSE3
:ID Q8CSE3 PRELIMINARY; PRT; 258 AA.
:AC Q8CSE3;
:DT 01-JUN-2001 (TReMBLrel. 17, Created)
:DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
:DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
:DE 2810411E22RIK protein (fragment).
:GN 2810411E22RIK
:OS Mus musculus (Mouse).
:OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
:OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
:OX NCBI_TaxID=10090;
:[1]
:SEQUENCE FROM N.A.
:STRAIN=C57BL/6J; TISSE=Embryo;
:MEDLINE=21085660; PubMed=11217851;
:RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehii Y.,
:RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
:RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
:RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
:RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
:RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
:RA Kuehl P., Lewis S., Mateuo Y., Nikolaic I., Pesole G., Quackenbush J.,
:RA Schirali L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
:RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
:RA Blake J., Boffelli D., Bojunga N., Carninci P., de Boraldo M. F.,
:RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
:RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,
:RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
:RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
:RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. F.,
:RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whittaker C., Wilming L.,
:RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
:RA Hayaishizaki Y.;
:RA "Functional annotation of a full-length mouse cDNA collection.";
:RT Nature 409:685-690(2001).
:DR EMBL; AK013082; BAB28637.1; -.
:DR MGD; MGI:1913563; 2810411E22RIK.
:DR InterPro; IPR000569; HECT_domain.
:DR Pfam; PF00632; HECT; 1.
:DR SMART; SM00119; HECTC; 1.
:DR PROSITE; PS50237; HECT; 1.
:FT NON_TER 1
:SEQUENCE 258 AA; 29670 MW; BOBE04BCB1PCBC98 CRC64;
```

```
Query Match          34.1%; Score 1378; DB 11; Length 258;
Best Local Similarity 98.8%; Pred. No. 2.7e-99;
Matches 255; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

> 491 GMAVPHGYIDGGTLPFFYKQLGKSIITDDMLVDPDLHNSLWILENDITGVLDHTFC 550
> 1 GMAVPHGYIDGGTLPFFYKQLGKSIITDDMLVDPDLHNSLWILENDITGVLDHTFC 60
>
> 551 VERNAYGEIIQHELKPKNGKSI PVNEENKKEYVLYNWRFLGIEAQFALOKGNEVIP 610
> 61 VERNAYGEIIQHELKPKNGKSI PVTEENKKEYVLYNWRFLGIEAQFALOKGNEVIP 120
>
> 611 QHLKLTDEKELELIICGLKIDVNDKVNRLKCHCTPDSNIVKFWKAVEFFDERRAR 670
> 121 QHLKLTDEKELELIICGLKIDVNDKVNRLKCHCTPDSNIVKFWKAVEFFDERRAR 180
>
> 671 LLQFVTGSSRVPLGGFKALOGAAGPRLFTTHQIDACTNNLPAHCTFNRIDIPPYESYEK 730
> 181 LLQFVTGSSRVPLGGFKALOGAAGPRLFTTHQIDACTNNLPAHCTFNRIDIPPYESYEK 240
>
> 731 LYEXLLTAIBETCGFAVE 748
> 241 LYEXLLTAIBETCGFAVE 258
>

RESULT 11
INT88      PRELIMINARY;      PRT;      820 AA.
> Q9NT88;
> 01-CCY-2000 (TrEMBLrel. 15, Created)
> 01-CCY-2000 (TrEMBLrel. 15, Last sequence update)
> 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
> Hypothetical protein (fragment).
> DKFZP434P2422.
> Homo sapiens (Human).
> Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
> Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
> NCBI_TaxID=9606;
> [1]
> SEQUENCE FROM N.A.
> TISSUE=Testis;
> Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
> Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
> -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
> EMBL; AL137469; CAB70754.1; -.
> HSSP; Q13526.1PIN.
> InterPro; IPR000008; C2.
> InterPro; IPR000569; HECT_domain.
> InterPro; IPR002349; WW.
> InterPro; IPR001202; WW_Rsp5_WWP.
> Pfam; PF00169; C2; 1.
> Pfam; PF00632; HECT; 1.
> Pfam; PF00397; WW; 3.
> PRINTS; PR00403; WWDOMAIN.
> SMART; SMO0239; C2; 1.
> SMART; SMO0119; HECTC; 1.
> SMART; SMO0456; WW; 3.
> SMART; SMO0456; WW; 3.
> PROSITE; PS0004; C2_DOMAIN_2; 1.
> PROSITE; PS0237; HECT; 1.
> PROSITE; PS0159; WW_DOMAIN_1; 3.
> PROSITE; PS0020; WW_DOMAIN_2; 3.
> Hypothetical protein.
> NON_TER
> SEQUENCE 820 AA; 95283 MW; 0FD334B29B3F4123 CRC64;

Query Match          33.8%; Score 1364.5; DB 4; Length 820;
Best Local Similarity 37.1%; Pred. No. 1.5e-97;
Matches 315; Conservative 115; Mismatches 221; Indels 199; Gaps 25;

> 50 TDTVNTLDPKRNQHYDLYIGKSD-SVTISVNNHKKHKQAGFICGVR----- 98
> 12 TKTKTKTLNPKNNEEFYRVNPSNHRLLPEVPDENLRTDD---FLGQVDVPLSHLPTD 68
```

```
QY 99 -----LLSNAINRLKDTGYQBLDLCKLGPNDNDTVRGQIVVUSLOSDEIGTGG 146
DB 69 PWERPYTKFLLAPRSHKRGVGLRLKWAYPKKG-----QCDEENSQORDMEHGW 123
QY 147 QVVCDSRLFPND-----LPDGWEERTASGRIQYLNHITRTTOWERPT--- 189
DB 124 EVDV-----SNDASQHOBELPPLPPPGWEEKVDNLGRTYVYVNNHNRITQWHRPSLMDV 178
QY 190 -----REASEYSSP-----CRPLSCFVDENTPISGTN-- 216
DB 179 SSEDNNTIRQINQEAHRRFRSRRHISEDLEPSEGGDVPEPWET-ISEEVNIAGSLG 237
QY 217 -----GATCGSSDPR-LAE---RRVR-----SORHNTYMSRTHL----- 247
DB 238 LALPPPPASPGSKTSPQELSELSRLQITPDSNGEQSSLIQREPPSSRLKSCVTDAVA 297
QY 248 ---HTPPD-----LPEGYEQRTTQGG 266
DB 298 EQGHLPPGAKDSFVRAVKDTLSNPQSPQSPYNSPKQHKVQSFLLPPGWEMRIAFNGR 357
QY 267 VYFLHTQTGTWTHDPRVP-----RDLNINCEBELGPLPGWEIRNTATGRVYFVDHNR 321
DB 358 PFIDENTTKTTTWEDPRLKFPVHVBKSTSLNPNDLGPLPGWEERIHLDCGTFYIDENSK 417
QY 322 TTQFTDPRLSANLHLVLRQNLKDQOQQQVVSLCPDDTECLTVPR--YKDLVQXKIL 379
DB 418 ITQWEDPRL-----QN-----PAITGPAVPYSRFPKQKYDYF 449
QY 380 ROELSOQQPQAGHCRIEYSRESRIPESYRVQVMKMPKDLWK-RLMKFRGEGLDYGVVA 438
DB 450 RKKLKPADIPNRFEMKLRNNIFESYRIRMSKRPDLVKARLWIEFESEKGLDYGVA 509
QY 439 REMWYLLSHMLNPPYGLFQYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMGMVPHG 497
DB 510 REMWFLLSKEMENPYGLFEXSATDNYTLQINPNSGLCNEDHLSYFTFGRVAGLAVPHG 569
QY 498 HYIDGGFTLPFYKQLGKSIITDDMLVDPDLHNSLWILENDITGVLDHTFCVHNAYG 557
DB 570 KLDDGFTIRFPYKMLGKQITLNDMESVDSYYSNKLWILENDFT-ELDLMFCIDENFG 628
QY 558 EIIQHELKPKNGKSI PVNEENKKEYVLYNWRFLGIEAQFALOKGNEVIPQHLKTF 617
DB 629 QTVQVDLKPENGSEIMVTNENKKEYIDLVIQNRFVNRVQKQNNAFLEGFTTELLPIDLIKIF 688
QY 618 DEKELELIICGLKIDVNDKVNRLK--HCTPDSNIVKFWKAVEFFDEERRARILQFV 675
DB 689 DENELELMCGLDVVDNDWRQHSIYKNGYC-PNHPVIQWFWKAVLLMDAEKRIQLDQFV 747
QY 676 TGSSRVPLQGFKALOGAAGPRLFTTHQIDACTNNLPAHCTFNRIDIPPYESYEKLYEKL 735
DB 748 TGTSRVPMNGFAELYSGNGPQLFTIEQWS-PEKLPAHCTCFNRLDLPPEYTFEDLREKL 806
QY 736 LTAIBETCGF 745
DB 807 LXAVENAQGF 816

RESULT 12
Q8BGE3      PRELIMINARY;      PRT;      887 AA.
ID Q8BGE3;
AC Q8BGE3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neural precursor cell expressed.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354693; PubMed=12466851;
```



```
SULT 14
2940
062940 PRELIMINARY; PRT; 887 AA.
01-NOV-1996 (TREMBlrel. 01, Created)
01-NOV-1996 (TREMBlrel. 01, Last sequence update)
01-MAR-2003 (TREMBlrel. 23, Last annotation update)
Nedd4 (Fragment)
NEDD4.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
MEDLINE=96221297; PubMed=8665844;
Staub O., Dho S., Henry P., Correa J., Ishikawa T., McGlade J.,
Rotin D.;
"WW domains of Nedd4 bind to the proline-rich PY motifs in the
epithelial Na+ channel deleted in Liddle's syndrome.";
EMBO J. 15:2371-2380(1996).
-|- SIMILARITY: CONTAINS 1 C2 DOMAIN.
EURL; U50842; AAB48949.1; -.
HSP; Q13526; 1PIN.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 3.
PRINTS; PR00360; C2DOMAIN.
SMART; SM00239; C2; 1.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 3.
PROSITE; PS00499; C2 DOMAIN 1; 1.
PROSITE; PS00004; C2 DOMAIN 2; 1.
PROSITE; PS50237; HECT; 1.
PROSITE; PS01159; WW DOMAIN 1; 3.
PROSITE; PS50020; WW DOMAIN 2; 3.
NON_TER 887 887
SEQUENCE 887 AA; D74B1097688CD9A1 CRC64;

Query Match
Best Local Similarity 33.6%; Score 1355.5; DB 11; Length 887;
Matches 314; Conservative 131; Mismatches 248; Indels 153; Gaps 25;

14 LRTVLCAKXVLKDFRLPDPFAKVV-----VDGSGQCHSTDTKNTLDPKWNQHYDLY- 68
77 VRVKVLAGLAKKDLGASDFVVRVTLVDPMSGVLTSVQTKIKSLNPKWNEI-LFR 135
69 -ICKSDSVISVWVHKKIKKQAGFLGVR-----LJLNAIN 105
136 VLPQQHRLIFVFDENLRFRD---FLGQVDVPLPLTENPRMERPYTFKDFVLHPRSH 192
106 RLKXDTGYQRDLCKLGFNDNDTVRGQIVVLSQSRDRIGTGQVVD---CSRL-----FDN 157
193 KSRVKGVLAKTYLPKNGSD-----ENADQAELEFGVWVLDQPDATHLOHPPPS 246
158 DLPDGHREARTASGRQYLNHTRTQMERPT-----RPAS 193
247 PLPFGWEERQDVLGRYYVYNHESRTTOMKRPSPEDDLTDENGDIQLQAHGAFTRQIS 306
194 E-----YSSPORPLSCFVDNTPISTNGATCGSSDP-----RLAE----- 230
307 EDVDPGDNHESPENWEIVREDNTIYSG-QAVQSPSPSGHPDVQVRLAEELDTRLTWGNP 365
231 ---RRVRSQRH---RNTMSRT-----HLHTPDLPEGYEQRTQQGQVYFLHTQT 274
366 ATSQPVTSSNHSRGGSSQTCIFEEQPTLPVLLPTSSGLPFGWEKQDRGRSYVVDHNS 425
275 GVSTWHPD-----RVP---RDLNINCELGPLPFGWEIRATGATGVYFVDHNRRT 323
```

```
Db 426 KTTWSKPTWQDDPSKIPAHLRGKTVPDSNDLGLPFGWEERTHTDGRVFFINHIKKT 485
Qy 324 QFTDRLSANLHLVLRNQQLKDOQQQWVSLCPDDTECLTVPRYKEDLVQKLIKLRQEL 383
Db 486 QWEDPRM-----QNVAITGPAEP-----YSRDYKRYEYFFRKL 519
Qy 384 SOQQAQAGHCRLEVSREEIFEESYRVQWVPEKDLWK-RLMIKRGEGGLDYGVAREWL 442
Db 520 KKTQDIPNKFEMKLRANILEDSTYRLMGVVRADFLKARLWIEFDGEGGLDYGVAREWF 579
Qy 443 YLLSHEMLNPYGLFQYS-RDDIYTLQINPDSAV-NPEHLSYHFVGRIMGMAVFFGHYI 500
Db 580 FLISKEMFNYPYGLFEYSATEDNYTLQINPNSGLCNEHLSYKFKIGRVAGMAYVHGKLL 639
Qy 501 DGGFTLPYKQLGKSIITLDDMELVDPDLNLSLWILENDITGVLDHTFCVEHNAVGEII 560
Db 640 DGGFTLPYKQLGKSIITLDDMELVDPDLNLSLWILENDITGVLDHTFCVEHNAVGEII 560
Qy 561 QHELKPNKSGSIPVNEENKKEYVRLVYVNRFLRGIEAQLALQKGFNEVIPHLLKTFDEK 620
Db 699 QHELKPGSEVVVTVNKNKKEYIYLVQWRFVNRIOQMAAFKSGFFELIPQDLIKIFDEN 758
Qy 621 ELELLICGLKIDVNDWKNTRLKH-CTPDSNIVKWPWKAVERFDEBERRARLLQFTVGSS 679
Db 759 ELELLMCGLDVNDWREHTKYKNGYSLAHQVHFWKAVLWMDSEKIRLLQFVTGTS 818
Qy 680 RVPLOQFKALQAGAPRLFTIHQIDACTNNLPKACTCNFRIDIPPYESYEKLYELLTAI 739
Db 819 RVPNMGFAELYGSGNGPSQFTVEQW-GTFDKLPRAHTCFNRLDLPPIESFDELNDKLQMAI 877
Qy 740 EETQGF 745
Db 878 ENTQGF 883

RESULT 15
Q9H2W4 PRELIMINARY; PRT; 854 AA.
AC Q9H2W4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE NEDD4La (Ubiquitin ligase NEDD4La).
GN NEDD4La OR NEDD4L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen H., McInnis M.G., DePaulo R. Jr., Ross C.A.;
RT "Identification of a novel gene on human chromosome 18q21 with
RT homology to NEDD4."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Chen H., Ross C.A., Wang N., DePaulo R. Jr., McInnis M.G.;
RT "NEDD4L on human chromosome 18q21 has multiple forms of transcripts
RT and is a novel member of the NEDD4 gene family."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210730; AAG43524.1; -.
DR EMBL; AF395931; AAM46208.1; -.
DR HSP; Q13526; 1PIN.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WNDOMAIN.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 4.
DR PROSITE; PS50020; WW DOMAIN 2; 4.
```

```

Ligase.
SEQUENCE      854 AA;  98181 MW;  00C74E161F52E7F CRC64;
Query Match      33.5%; Score 1354.5; DB 4; Length 854;
Best Local Similarity 43.7%; Pred. No. 9.9e-97;
Matches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;

137 QSRDRIGTGQVVDCSRUFDN---DLPGWEERTASGRIOYLNHIHITTCWERPRPAS 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 RARSSTVTCGEEPTSAVYHTTGPLSGWEERKDAKRTYYVNNHNRRTTWTTPINOLA 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
194 EYSSPG-----RPLSC---FVDENTPISGINGATCGOSSDPRLAEREVRS 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 EDGASGATNSNNHLEIPOIRPRSLSPVTLSAPLEGAKDSPVRAVXDTLSNPQSQ 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
236 QRHRNYSRTHLHTPDLPBEGYEQRTTQGGVYFHTGTGVSTWHDPRVP-----RDLN 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
361 PSFYSKPKQHKVQTSFPPGWMEMRIANGRPFFIDHNKTKTTWEDPRLKFPVHMRSKTS 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
291 INCEELGPLPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRQNLKDOQQQ 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
421 LNENDLGPLPGWEERIHLDTGRTFYDHSKITQWEDPRL-----QN----- 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
351 QVVSICPDTECLTVPR--YKDLVQKLIKRLQELSQOQPOAGHCRIEVSREBIFESYR 408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
463 -----PAITGPANVPYSREFKQKYDFRKKLKKPADIPNRFENKLRNNIFEESYR 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409 QVMKMRPKDLWK-RLMIKPRGEGLDYGVAREWLYLLSHEMLNPPYGLPOYSRDDIYTL 467
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
513 RINSVKRPDLKARLWIEFESEKGLDYGVAREWFFLLSKEMFNPPYGLFEYSATDNYTL 572
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
468 QINPDSAV-NPEHLSYFHFVGRIMGMVFGHYIDGGFTLFFYKQLLGKSIITLDDMELVD 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
573 QINPNSGLCNEDHLSYFTFGRVAGLAVFHGKLLDGFIRFPFYKMLGKQITLNDMESVD 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
527 PDLHNSLWILENDITGLDHTFCVEHNAYGEILIOHELKPNKSIIPVNEKNKEYVELYV 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
633 SEFYNSLKWILENDPT-ELDLNFCIDENFGTYQVVDLKNPNSSEIMVTNENKREYIDLVI 691
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
587 NNRFLRGIEAQLAQGFNEVIPQHLKTFDEKELELIICGLGKIDVNDKVNTRLK-- 644
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
692 QNRFVNRVQKQNAFLEGTELLPIDLIKIFDENELELLMCGLDGVDVNDWRQHSIYKNG 751
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
645 HCTPDSNIYKWKAVEFDEERRARILQFVTGSSRPVLOGFKALQAGAPRLFTIHQID 704
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
752 YC-FNHPVIOFWFKAVLLMDAEKRIRLQFVTGTSRVPMPNGFAELYGNGPQLFTIEQWG 810
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
705 ACTNNLPKAHTCFNRIDIPPPYESYEKYLITAEETCGF 745
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
811 S-PEKLPRAHTCFNRLDLPPTFTEDLREKULMAVENAQGP 850
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

arch completed: February 20, 2004, 15:29:57
b time : 39.0863 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

M protein - protein search, using sw model

un on: February 20, 2004, 15:25:26 ; Search time 11.1869 Seconds
(without alignments)
3144.377 Million cell updates/sec

itle: US-10-009-945-4

effect score: 4038

equences: 1 MSNPGRRNGPVKRLTLVLC.....EKLYKLLTAIBETCGPAVE 748

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 127863 seqs, 47026705 residues

total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atadase : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	4031	99.8	748	1 SUP2 HUMAN	Q9hau4 homo sapien
2	3015.5	74.7	757	1 SUP1 HUMAN	Q9hce7 homo sapien
3	2937.5	72.7	731	1 SUP1 XENLA	Q9pun2 xenopus lae
4	2491.5	61.7	619	1 SUP1 MOUSE	Q9cun6 mus musculus
5	1585.5	39.3	767	1 PUB1 SCHPO	Q92462 saccharomyc
6	1484	36.8	809	1 RFP5 YEAST	P39940 saccharomyc
7	1358	33.6	957	1 NED4 MOUSE	P46935 mus musculus
8	1354.5	33.5	927	1 NED4 HUMAN	P46934 homo sapien
9	626	15.5	310	1 UR81 RAT	P51593 rattus norv
10	539	13.3	875	1 UR3A HUMAN	Q05086 homo sapien
11	484.5	12.0	885	1 UR3A MOUSE	Q08759 mus musculus
12	460.5	11.4	892	1 HUL4 YEAST	P40985 saccharomyc
13	445.5	11.0	1050	1 HER3 HUMAN	Q15034 homo sapien
14	404	10.0	1992	1 TRIB HUMAN	Q14669 homo sapien
15	342	8.5	910	1 HUL5 YEAST	P53119 saccharomyc
16	286.5	7.1	1620	1 HED1 HUMAN	Q9ult8 homo sapien
17	281	7.0	920	1 EDD RAT	Q82671 rattus norv
18	280	6.9	2799	1 EDD HUMAN	Q95071 homo sapien
19	278.5	6.9	1483	1 YF4 YEAST	P33202 saccharomyc
20	262.5	6.5	1647	1 YF31 SCHPO	Q10435 schizosacch
21	224.5	5.6	2895	1 HYD DROME	P51592 drosophila
22	207	5.1	472	1 YAP1 MOUSE	P46938 mus musculus
23	158	3.9	448	1 YAP1 CHICK	P46936 gallus gall
24	154.5	3.8	454	1 YAP1 HUMAN	P46937 homo sapien
25	122.5	3.0	672	1 KP4A HUMAN	P17252 homo sapien
26	122.5	3.0	672	1 KP4A RABIT	P10102 oryctolagus
27	122.5	3.0	672	1 KP4A RAT	P05696 rattus norv
28	122.5	3.0	637	1 UN13 YEAST	Q08484 saccharomyc
29	120.5	3.0	1813	1 UN13 CAEBL	P27715 caenorhabdi
30	119.5	3.0	593	1 CNE5 HUMAN	Q9hch3 homo sapien
31	116.5	2.9	583	1 PR40 YEAST	P33203 saccharomyc
32	116.5	2.9	672	1 KP4A MOUSE	P20444 mus musculus
33	116	2.9	163	1 PIN1 HUMAN	Q13526 homo sapien

ALIGNMENTS

RESULT 1

ID	SUP2 HUMAN	STANDARD;	PRT;	748 AA.
AC	Q9HAU4: Q9H260;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Smad ubiquitination regulatory factor 2 (EC 6.3.2.-) (Ubiquitin--			
DE	protein ligase SMURF2) (Smad-specific E3 ubiquitin ligase 2)			
DE	(hSMURF2).			
GN	SMURF2			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF PRO-251--VAL-284 AND			
RP	GLY-297--LEU-330.			
RX	PubMed=1163210;			
RA	Kaveak P., Rasmussen R.K., Causing C.G., Bonni S., Zhu H.,			
RA	Thomsen G.H., Wrana J.L.;			
RT	"Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the			
RT	TGF-beta receptor for degradation.";			
RL	Mol. Cell 6:1365-1375(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF PRO-251--VAL-284 AND CYS-716.			
RP	MEDLINE=20538422; PubMed=11016919;			
RX	Lin X., Liang M., Feng X.-H.;			
RT	"Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent			
RT	degradation of Smad2 in transforming growth factor-beta signaling.";			
RL	J. Biol. Chem. 275:36818-36822(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-716.			
RX	MEDLINE=21107656; PubMed=11158580;			
RA	Zhang Y., Chang C., Gehling D.J., Hemmati-Briylanlou A., Derynck R.;			
RT	"Regulation of Smad degradation and activity by Smurf2, an E3			
RT	ubiquitin ligase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:974-979(2001).			
CC	-!- FUNCTION: Interacts with SMAD1, SMAD2 and SMAD7 in order to			
CC	trigger their ubiquitination and proteasome-dependent degradation.			
CC	Enhances the inhibitory activity of SMAD7 and reduces the			
CC	transcriptional activity of SMAD2. Coexpression of SMURF2 with			
CC	SMAD1 results in considerable decrease in steady-state level of			
CC	SMAD1 protein and a smaller decrease of SMAD2 level.			
CC	-!- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, SMAD6 and SMAD7 but			
CC	not SMAD4.			
CC	-!- SUBCELLULAR LOCATION: Nuclear. Cytoplasmic in the presence of			
CC	SMAD7.			
CC	-!- TISSUE SPECIFICITY: Widely expressed.			
CC	-!- DOMAIN: The second and third WW domains are responsible for			
CC	interaction with R-SMAD (SMAD1, SMAD2 and SMAD3).			
CC	-!- SIMILARITY: Contains 1 C2 domain.			
CC	-!- SIMILARITY: Contains 3 WW domains.			
CC	-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase			
CC	domain.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
EMBL; AF310676; AAG45422.1; -
EMBL; AF301463; AAG25641.1; -
EMBL; AY014180; AAG50421.1; -
HSP; Q13526; 1PIN.
MIM; 605532; -.
GO; GO:0004842; F:ubiquitin-protein ligase activity; NAS.
GO; GO:0016481; P:negative regulation of transcription; NAS.
GO; GO:0017015; P:regulation of TGFbeta receptor signaling pa...; NAS.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF06632; HECT; 1.
Pfam; PF00397; WW; 3.
SMART; SM00239; C2; 1.
SMART; SM00119; WW; 3.
SMART; SM00112; HECTC; 1.
SMART; SM00456; WW; 3.
PROSITE; PS00499; C2 DOMAIN_1; 1.
PROSITE; PS00504; C2 DOMAIN_2; 1.
PROSITE; PS02037; HECT; 1.
PROSITE; PS01159; WW DOMAIN_1; 1.
PROSITE; PS00200; WW DOMAIN_2; 3.
Ubl conjugation pathway; Ligase; Repeat; Nuclear protein.
DOMAIN 1 98 C2 DOMAIN.
DOMAIN 157 190 WW 1.
DOMAIN 251 284 WW 2.
DOMAIN 297 330 WW 3.
DOMAIN 414 748 HECT.
MUTAGEN 251 284 MISSING: ABOLISHES INTERACTION WITH SMAD2 AND SMAD7.
MUTAGEN 297 330 MISSING: ABOLISHES INTERACTION WITH SMAD7.
MUTAGEN 716 716 C-A: LOSS OF ABILITY TO UBIQUITINATE SMAD2.
MUTAGEN 716 716 C-G: LOSS OF ACTIVITY (LOSS OF ABILITY TO UBIQUITINATE SMAD1 AND SMAD2 AND NO DOWN-REGULATION OF SMAD1 AND SMAD2 PROTEIN LEVELS).
CONFLICT 6 6 G -> R (IN REF. 2).
SEQUENCE 748 AA; 86195 MW; 3042B443A3755762 CRC64;
```

Query Match 99.8%; Score 4031; DB 1; Length 748;
Best Local Similarity 99.9%; Pred. No. 1.6e-280;
Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MSNPGRRNGVKRLVLCANLVKKDFRLPDPFAKVVDSGGCHSTDTVQNTLDPK 60
1 MSNPGRRNGVKRLVLCANLVKKDFRLPDPFAKVVDSGGCHSTDTVQNTLDPK 60
61 WQHYDLYIGKSDVTSVWNHKKHKKQAGFLGCVRLLSNAINLKDGTGYRLDCLK 120
61 WQHYDLYIGKSDVTSVWNHKKHKKQAGFLGCVRLLSNAINLKDGTGYRLDCLK 120
121 GPNDNDTVRGQIVVVSLOSRRIGTGQGVDCSLFNDLPDQWEERTASGRIQYLNHIT 180
121 GPNDNDTVRGQIVVVSLOSRRIGTGQGVDCSLFNDLPDQWEERTASGRIQYLNHIT 180
181 ETTQWERPTRASSYSPGRPLSCFVDENPTISGTGATCGOSSDRLAERVRVSRHRN 240
181 ETTQWERPTRASSYSPGRPLSCFVDENPTISGTGATCGOSSDRLAERVRVSRHRN 240
241 YMSRTHLTPDLPDEGYEQRTTQGGVYFLHTGTGSTVHDPVRPDLNSNINCEELGFLP 300
241 YMSRTHLTPDLPDEGYEQRTTQGGVYFLHTGTGSTVHDPVRPDLNSNINCEELGFLP 300

```
Qy 301 PGWEINTATGRVYFDVHNRTTQTDPRLSANLHLVLRQNLKDQOQOQVWSLCFDDT 360
Db 301 PGWEINTATGRVYFDVHNRTTQTDPRLSANLHLVLRQNLKDQOQOQVWSLCFDDT 360
Qy 361 ECLTVPRYKRDVQKLKILRQELSQQQPAQGHCRLEVSREEIFFEESYRQVMKMPKOLWK 420
Db 361 ECLTVPRYKRDVQKLKILRQELSQQQPAQGHCRLEVSREEIFFEESYRQVMKMPKOLWK 420
Qy 421 RLMIKFRGEGLDYGVAVREWLILSHEMLNPNPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Db 421 RLMIKFRGEGLDYGVAVREWLILSHEMLNPNPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Qy 481 SYEHFVGRIMGVAFHGHVYIDGGFTLPVFKQLLGSITLDDMELVDPDLHNSLVWILEND 540
Db 481 SYEHFVGRIMGVAFHGHVYIDGGFTLPVFKQLLGSITLDDMELVDPDLHNSLVWILEND 540
Qy 541 ITGVLDHTFCVEHNAYGEIIQHELKPNKGSIPVNEENKKEYVRLYVNWRFJRGIEAQFLA 600
Db 541 ITGVLDHTFCVEHNAYGEIIQHELKPNKGSIPVNEENKKEYVRLYVNWRFJRGIEAQFLA 600
Qy 601 LQKGFNEVPOHLLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660
Db 601 LQKGFNEVPOHLLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660
Qy 661 EFDEERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQIDACTNNLPKATCFNRI 720
Db 661 EFDEERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQIDACTNNLPKATCFNRI 720
Qy 721 DIPVESYKLYEKLITAEETCGFAVE 748
Db 721 DIPVESYKLYEKLITAEETCGFAVE 748

RESULT 2
SUFI_HUMAN STANDARD; PRT; 757 AA.
AC QHCE7; O75853; Q9UUT8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
(hSMURF1).
DE SMURF1 OR KIAA1625.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 10-731 FROM N.A. (ISOFORM SHORT).
RX MEDLINE=9935348; PubMed=10458166;
RT Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RA "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
pattern formation."
RL Nature 400:687-693(1999).
RN [2]
RP SEQUENCE OF 20-731 FROM N.A. (ISOFORM SHORT).
RX Stoneking T., Bauer C., O'Neal D.;
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Brain;
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII: the complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
CC -I- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION.
CC -I- ALTERNATIVE PRODUCTS.
CC Event=Alternative splicing; Named isoforms=2;
```

```
Names-Long;
Isid=Q9HCE7-1; Sequence=Displayed;
Names-Short;
Isid=Q9HCE7-2; Sequence=VSP_006812;
-|- SIMILARITY: Contains 1 C2 domain.
-|- SIMILARITY: Contains 2 WW domains.
-|- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF199364; AAF08298.2; --
EMBL; AC004893; AAC62434.1; --
EMBL; AB046845; BAB13451.1; ALT_INIT.
HSP; Q13526; IPIN.
MIN; 605568; --
GO; GO:0005622; C:intracellular; TAS.
GO; GO:0000211; F:protein degradation tagging activity; IDA.
GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
GO; GO:0030154; P:cell differentiation; IDA.
GO; GO:0007398; P:ectoderm development; TAS.
GO; GO:0030514; P:negative regulation of BMP signaling pathway; TAS.
GO; GO:0006464; P:protein modification; TAS.
GO; GO:0006512; P:ubiquitin cycle; IDA.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.
InterPro; IPR01202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 2.
SMART; SM00239; C2; 1.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 2.
PROSITE; PS00499; C2 DOMAIN 1; 1.
PROSITE; PS00094; C2 DOMAIN 2; 1.
PROSITE; PS0237; HECT; 1.
PROSITE; PS01159; WW DOMAIN 1; 1.
PROSITE; PS00020; WW DOMAIN 2; 2.
Ub1 conjugation pathway; Ligase; Repeat; Alternative splicing.
DOMAIN 1 99 C2 DOMAIN.
DOMAIN 234 267 WW 1.
DOMAIN 306 339 WW 2.
DOMAIN 420 757 HECT.
BINDING 725 725 UBIQUITIN.
VARSPIC 269 294 Missing (in isoform Short).
/FTId=VSP_006812.
MUTAGEN 725 725 C->A; LOSS OF UBIQUITINATION CAPACITY.
SEQUENCE 757 AA; 86113 MW; 89A171CFC47B40E9 CRC64;

Query Match
Best Local Similarity 74.7%; Score 3015.5; DB 1; Length 757;
Matches 574; Conservative 63; Mismatches 73; Indels 85; Gaps 10;

1 MSNPFRRNG-PVKRLTLVLCANLVKDPFPLPDPFAKVVDSGGQCHSDTVKVLDP 59
1 MSNPFRRNGSSIKRLTLVLCANLVKDPFPLPDPFAKVVDSGGQCHSDTVKVLDP 60
60 KMNQHYDLVYKSDSVTISVWNHKKHKQAGFLGCVRLLSNAINRLKDTGYQLDLCK 119
61 KMNQHYDLVYKSDSVTISVWNHKKHKQAGFLGCVRLLSNAINRLKDTGYQLDLCK 120
120 LGNDNDTVRGQIVWSLQSDRIGTGQGVQVDCSLFNDLDPGWEERTASGRQYLNHI 179
121 LNPSTDVARGQIVWSLQSDRIGTGQGVQVDCSLFNDLDPGWEERTASGRQYLNHI 163
180 TRTQWERPRTPASEYSSGRPLSCFVDENVTISGTNGATCG-----QSSDPLAE 230
```

```
Db 164 -----EDSGRGRPLSCFMEEAPYTDSTCAAGAGGNCRFVSPSQDORLQA 209
QY 231 RVRVSQRHNYM-----SRTHLTPDPLPEGVEQSTTQGGQVYFLHTOTGVSTWHDPRV-- 284
Db 210 QRLRNPDRVGRSLQTPQNRPHGQSPFEGVEQSTTQGGQVYFLHTOTGVSTWHDPRIPS 269
QY 285 -----PRDLSNINCEELGFLPPGWEIRNATGRVYFVDHNN 320
Db 270 PSGTIPGGDAALYBFLQCHTSFRLNSVNCDELGLPFGWEVRSTVSGRIYFVDHNN 329
QY 321 RTQTFDRLSLNHLVLRNQKLDKQKQKQV-----SLCPDDTECLTVPRYKRLVQKL 376
Db 330 RTQTFDRL-----LHIMHQCLKPSQPLPSEGL-----EDEELPAQRYERDLVQKL 382
QY 377 KILROELSQOQAGCHRIEVSREIEPESYRQVWKRPDKLWKRLMKRGEGLDYG 436
Db 383 KVLRELSLQOQAGCHRIEVSREIEPESYRQVWKRPDKLWKRLMKRGEGLDYG 442
QY 437 VAREWLYLLSHEMLNPPYGLFYQSRDDIYTLQINPDSAVNPEHLSYFHFVGRIMGMAVFF 496
Db 443 VAREWLYLLCHEMLNPPYGLFYQSRDDIYTLQINPDSAVNPEHLSYFHFVGRIMGMAVFF 502
QY 497 GHVIDGGFTLPEYKOLLGKSLTLDMLVDPDLNLSLVILENDITGVLDHTFCVENAY 556
Db 503 GHYINGGFTVPYKOLLGKPIQLSDLESVDPELHLSLVILENDITGVLDHTFCVENAF 562
QY 557 GRIIQLHLPKNGKSIIPVNEENKKEYRVLVYNWRPLRGIEAQFALQKGFNEVIPHLLKT 616
Db 563 GRILQHELKNGRNVPTVEENKKEYRVLVYNWRPLRGIEAQFALQKGFNEVIPHLLKP 622
QY 617 FDEKELELIIIGLKDIDNDKNSNRLKCHCTPDSNIVKFWKAVEFDEERRARLLQFVT 676
Db 623 FDEKELELIIIGLKDIDNDKNSNRLKCHCTPDSNIVKFWKAVEFDEERRARLLQFVT 682
QY 677 GSSRVPLQGFALQ-----GAAGFRLLTHIDACTNNLPKHTCENRIDIPYSEYKLYE 733
Db 683 GSTRVPLQGFALQSGTGAAGFRLLTHIDACTNNLPKHTCENRIDIPYSEYKLYE 742
QY 734 KLLTAIBETCGFAVE 748
Db 743 KLLTAIBETCGFAVE 757

RESULT 3
ID SCFL_XENLA STANDARD; PRT; 731 AA.
AC Q9PUN2;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 1 (SC 6.3.2.-) (Ubiquitin--
protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase) (xSMURF1).
GN SMURF1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blastula;
RX MEDLINE=99385348; PubMed=10458166;
RA Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
RL pattern formation."
CC Nature 400:687-693(1999).
CC -|- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION. MAY
CC REGULATE ECTODERMAL DIFFERENTIATION AND PATTERN BY MODULATING BMP
CC SIGNALING AND MAY ENHANCE CELLULAR RESPONSIVENESS TO THE SMAD2
CC (ACTIVIN/TGF-BETA) PATHWAY.
CC -|- DEVELOPMENTAL STAGE: EXPRESSED FROM THE EGG STAGE TO THE SWIMMING
```

200 DONINCERBUSEJFOWEIKNIMIGRAJZVDHNNKTIYQZTDFKJUSANUHWVUNAKQNUKZKQ

CC -I- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR

0001

Db	56	WNETFEVNTVNDSTIAIQVFDQKFKKKQGGLGVINLAVGVDLDAIGGDEMLRDLK	114
Qy	119	KLGPNDNDVRGQITWISLQ	137
Db	115	K--SNENTVTVHGKIIINLSTTAQSTLQVPSAASGARTQRTSITNDPQSQSSSVSENPA	172
Qy	138	-----SRDRITGGQVVCDSRLFDN-----DLPDGEERTAGRTQYLMHITRTIQ	184
Db	173	SSRAGSPTRDNAPAAASPASSEPRTFSSFEDQYGLPPGWERRTDNLGRVTVYVDHNRSTT	232
Qy	185	WERPT-----RPASEYSSPGRLPSCFVDSENTPISTNGATCGSSDPRLABRRVRSQRH	238
Db	233	WIRPILSVAGAAAELHSSAS--SANVTGVPSSNAA-----RRTASVLIT--	279
Qy	239	RNYMSRTHLTPDLPPEGYEQRTOOQGVYFLHTQTGVSTWHDPRVPRDLNIN-----	292
Db	280	---SNATTAGSGELPGWEQRVTPGPRPYFVDHNRTRTTTWDPFRQVIRSYGGPNNA	335
Qy	293	-----CBELGPLPGWIRNTAGRVYFVDHNRRTQFTDPRLSANLHLVFNQQLKQ	347
Db	336	IQQPVSGQLPGSGWEMLNTARVYFVDHNRKTTTWDPRPSSLS	384
Qy	348	QQQVSVLCPDDTECLTVPRYKRLVOKLILRQELSSQ--QPQAGHCRIEVSREEIFEE	405
Db	385	-----NVPQYKRDFRKLIYF--LSQPALHPLPGQCHIKVERNHI	425
Qy	406	SYRQVMKMRPKDILMKRLMKRBEGLDYGGAWEMLYLLSHMLNPYGLFOYSRDDIY	465
Db	426	SYAIRMQSATDLKKRLMKLXFDGEDGJYGLGREYFLLSHSMFNPFCYLFYSSVDNY	485
Qy	466	TLQNPDSAVNPEHLSYFHFVGRIMGMAVFHGHIQDGFPLPYKOLLGKSIITLDDNELV	525
Db	486	TLQINPHSGINPEHLNFKYIGRIVGIGLAIETHRFVDAFPVVSFYKMLQKVTQLQDMES	545
Qy	526	DPDLHNSLVITLNDITGVLDHTFCVEHNAYGIIIOHELKPNKGSIPVMEENKKEYVRLY	585
Db	546	DAEYRSVLWTLDNDITGVLDITFSVEDNCFGSVVITDLKPNGRNIEVTEENKREYVDLV	605
Qy	586	VNWAFRLGIEAQFALOKGFNEVPIQOHLKATFDEKELELIICGLGKIDVNDKVNTRLKH	645
Db	606	TWVLIQRIEIQFNAFHEGSELIQBELINVPDERELLELIIGISIDMEDWKHTDYRS	665
Qy	646	CTPDSNIVKFWKAVEFFDEBRARLLQFVTGSSRVPLOFGKALQAGAPRFTTHQIDA	705
Db	666	YSENDQIIKFWELMDWSNEKKSRLLQFTTGTSRIPVANGFKDQSGDGRPKRTIEKAGE	725
Qy	706	CTNNLPKATCFNEDIDIPPYESYEKYLKTLAETCCFAVE	748
Db	726	-PNKLUPKATCFNEDLPPYTSKKDLQHLKLSIAVEETIGFQGE	767
RESULT 6			
RSP5 YEAST			
ID	RSP5 YEAST	STANDARD;	PRT; 809 AA.
AC	R39340;		
DC	01-FEB-1995 (Rel. 31, Created)		
DT	01-FEB-1995 (Rel. 31, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Ubiquitin-protein ligase RSP5 (EC 6.3.2.-)		
GN	RSP5 OR NP11 OR MDPI OR VER125W OR SYGP-ORF41.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
ON	NCBI_TaxID=4932;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=S288C / A3972;		
RY	PubMed=919868;		
RA	Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,		
RA	Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,		
RA	Cherry J.M., Chung E., Duncan M., Guzman B., Hartzell G.,		
RA	Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,		

Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C., Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
"The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V."; Nature 387:78-81(1997).
[2] IDENTIFICATION.
Winston F.;
Unpublished observations (FEB-1993).
[3]
CHARACTERIZATION.
STRAIN=Sigma 1278B;
MEDLINE=96154942; PubMed=8596462;
Hein C., Springsael J.-Y., Volland C., Haguenaer-Tsapis R., Andre B.;
"NPII, an essential yeast gene involved in induced degradation of Gsp1 and Fur4 permeases, encodes the Rsp5 ubiquitin-protein ligase."; Mol. Microbiol. 18:77-87(1995).
[4]
FUNCTION.
MEDLINE=95223981; PubMed=7708685;
Huibregtse J.M., Scheffner M., Beaudenon S., Howley P.M.;
"A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase."; Proc. Natl. Acad. Sci. U.S.A. 92:2563-2567(1995).
[5]
ERRATUM.
MEDLINE=95281634; PubMed=7761480;
Huibregtse J.M., Scheffner M., Beaudenon S., Howley P.M.;
Proc. Natl. Acad. Sci. U.S.A. 92:5249-5249(1995).
[6]
CHARACTERIZATION.
STRAIN=PV56;
MEDLINE=99077972; PubMed=98589558;
Wang G., Yang J., Huibregtse J.M.;
"Functional domains of the *rsp5* ubiquitin-protein ligase."; Mol. Cell. Biol. 19:342-352(1999).
-!- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES. REQUIRED FOR UBIQUITINATION AND THEREFORE DEGRADATION OF SEVERAL CELL SURFACE PROTEINS LIKE GAP1, FUR4, MAL61 AND STE2. ALSO ACTS ON RBP1.
-!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (Potential).
-!- PTM: THE UBIQUITINATION APPEARS TO BE THE RESULT OF AN INTRAMOLECULAR TRANSFER OF UBIQUITIN.
-!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-thioester formation.
-!- SIMILARITY: Contains 1 C2 domain.
-!- SIMILARITY: Contains 3 WD domains.
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.
-!- SIMILARITY: STRONG, TO S.POMBE PUB1.
-!- SIMILARITY: TO YEAST YKL010C.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcements> or send an email to license@isb-sib.ch).

EMBL; U18916; AAC03223.1; -.
FIR; S43217; S43217.
HSP; Q13526; 1PIN.
SGD; S000927; RSP5.
GO; GO:000151; C:ubiquitin ligase complex; IDA.
GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
GO; GO:0006333; P:chromatin assembly/disassembly; IMP.
GO; GO:0006513; P:protein monoubiquitination; IDA.
GO; GO:0000209; F:protein polyubiquitination; IDA.
InterPro; IPR000008; C2.
InterPro; IPR000569; HECT domain.

677 KLLDGFIRFFYKMLHKEPITLHDMESVDSEYNSLRWILENDPT-ELDLRFIIDEELPG 735
558 EIIQHELKNGKSIPIVNEENKKEYVRLVNRFLRGIEAQLALQKGFNEVIPHLLKTFP 617
736 OTHQHELKNGGSEIVVTKNKKKEYTYLVQWRVNRRIQKQMAAFKEGFFELIPDOLKIF 795
618 DEKELELIICGLGKIDVNDKVNTELKH-CTPDSNIVKFWKAVEFFDEERARLLQFVT 676
796 DENELELLMCGLDVNDVNDREHTKYKGYNSANHOVQVFWKAVLMDSEKRIELLOFVT 855
677 GSRVLPQKALQGAAGPRLFTIHQIDACTNNLPKAHTCFNRIDIPPYSEYKLYSKLL 736
856 GTSRVPVNGFAELYGNGPSQSTVBQW-GTPEKLPRAHTCFNRDLDPYSEFELWDLQ 914
737 TAJEETCGF 745
915 MALENTQGF 923

SULT 9
BI_RAT STANDARD; PRT; 310 AA.
P51593;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
DNA binding protein URE-B1 (EC 6.3.2.-).
UREB1.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
TISSUE=Brain;
MEDLINE=95058008; PubMed=7968380;
Gu J., Ren K., Dubner R., Iadarola M.J.;
"Cloning of a DNA binding protein that is a tyrosine kinase substrate
and recognizes an upstream initiator-like sequence in the promoter of
the preprodynorphin gene."
Brain Res. Mol. Brain Res. 24:77-88(1994).
-!- FUNCTION: MAY BE A DNA-BINDING TRANSCRIPTIONAL REGULATOR.
RECOGNIZES AN UPSTREAM INITIATOR-LIKE SEQUENCE IN THE PROMOTER OF
THE PREPRODYNORPHIN GENE.
-!- FUNCTION: E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS UBIQUITIN FROM
AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A THIOESTER AND
THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED SUBSTRATES (BY
SIMILARITY).
-!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
thioester formation.
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U08214; AAA81950.1; -
InterPro; IPR000569; HECT_domain.
Pfam; PF00632; HECT; 1.
SMART; SM00119; HECTC; 1.
PROSITE; PS0237; HECT; 1.
UBI conjugation pathway; Ligase; Repeat; DNA-binding;
Transcription regulation.
DOMAIN 202 310 HECT.
BINDING 275 275 UBIQUITIN.
SEQUENCE 310 AA; 35640 MW; D62E7F21F47C068E CRC64;
Query Match 15.5%; Score 626; DB 1; Length 310;

Best Local Similarity 44.1%; Pred. No. 1.5e-37;
Matches 135; Conservative
Indels 2; Gaps 2;
444 LLSHMLNPPYGLFOYSDDDIYTLQNPDSAVNPEHLSVHFVGRIMGVAVPHGHVLDGG 503
2 IISREMFNYALFRTSPGVRVTRSHLPTVPTTSSSLAGRYVAKAVMTALLECY 61
504 FTLFPYKOLLGKSIITLDDMELVDPDLHNSLVWILENDITGV-LDHTFCVHNAYGEIIQH 562
52 FTRSPYKHILGKSVRYTDMSESDVHFYQGLVYLENDVSTGLDYDLTFSVEVQFGVCEVR 121
563 ELKNGKSIPIVNEENKKEYVRLVNRFLRGIEAQLALQKGFNEVIPHLLKTFPEKEL 622
122 DLKNGANILVTEENKKEYVRLVNRFLRGIEAQLALQKGFNEVIPHLLKTFPEKEL 181
623 ELIICGLGKIDVNDKVNTELKH-CTPDSNIVKFWKAVEFFDEERARLLQFVTGSSRVP 682
182 ELLYTGLEFTIDDLKNTVEHYKQNSIQVFWRALRSFDOADRAKFLQFVTGTSKVP 241
683 LQGFALQGAAGPRLFTIHQIDACTNNLPKAHTCFNRIDIPPYSEYKLYSKLLTAEET 742
242 LQGFALQGAAGPRLFTIHQIDACTNNLPKAHTCFNRIDIPPYSEYKLYSKLLTAEET 301
743 C-GFAY 747
302 SEGFG 307
RESULT 10
UE3A HUMAN STANDARD; PRT; 875 AA.
ID UE3A HUMAN STANDARD; PRT; 875 AA.
AC Q05086; P78355; Q93066;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3A (EC 6.3.2.-) (Oncogenic protein-
associated protein E6-AP) (Human papillomavirus E6-associated
protein).
DE UBE3A OR E6AP.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM I; II AND III).
RC TISSUE=Keratinocytes;
RX MEDLINE=9728825; PubMed=9143503;
RA Yamanoto Y., Huijbreghse J.M., Howley P.M.;
"The human E6-AP gene (UBE3A) encodes three potential protein isoforms
generated by differential splicing."
RL Genomics 41:263-266(1997).
RN [2]
RP SEQUENCE OF 11-875 FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Keratinocytes;
RX MEDLINE=93140775; PubMed=8380895;
RA Huijbreghse J.M., Scheffner M., Howley P.M.;
"Cloning and expression of the cDNA for E6-AP, a protein that
mediates the interaction of the human papillomavirus E6 oncoprotein
with p53."
RT Mol. Cell. Biol. 13:775-784(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM I), AND VARIANT VAL-290.
RC TISSUE=Fetal brain;
RX MEDLINE=97141921; PubMed=8988171;
RA Kishino T., Lalonde M., Wagstaff J.;
"UBE3A/E6-AP mutations cause Angelman syndrome."
RL Nat. Genet. 15:70-73(1997).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM I), AND VARIANTS TYR-44; THR-201 AND
VAL-290.
RX MEDLINE=97141922; PubMed=8988172;
RA Matsura T., Sutcliffe J.S., Fang P., Galjaard R.-J., Jiang Y.-H.,
Benton C.S., Rommens J.M., Beaudet A.L.;

"De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome.";
Nat. Genet. 15:74-77(1997).
[5]
VARIANT AS ILE-826 INS, AND VARIANTS HIS-62; THR-201 AND PRO-372.
MEDLINE=98254455; PubMed=9585605;
Maizac P., Webber H., Moncla A., Graham J.M. Jr., Kukulich M.,
Williams C., Pagon R.A., Ramesh L.A., Kishino T., Wagstaff J.,
"Mutation analysis of UBE3A in Angelman syndrome patients.";
Am. J. Hum. Genet. 62:1353-1360(1998).
[6]
CHARACTERIZATION.
MEDLINE=983151563; PubMed=9688277;
Nuber U., Schwarz S.E., Scheffner M.;
"The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as
its own substrate";
Eur. J. Biochem. 254:643-649(1998).
-!- FUNCTION: INTERACTS WITH THE E6 PROTEIN OF THE CANCER-ASSOCIATED
HUMAN PAPILLOMAVIRUS TYPES 16 AND 18. THE E6/E6-AP COMPLEX BINDS
TO AND TARGETS THE P53 TUMOR-SUPPRESSOR PROTEIN FOR UBIQUITIN-
MEDIATED PROTEOLYSIS. IT IS AN E3 UBIQUITIN-PROTEIN LIGASE WHICH
ACCEPTS UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE
FORM OF A THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO
TARGETED SUBSTRATES. IT CAN TARGET ITSELF FOR UBIQUITINATION IN
VITRO AND EFFICIENTLY PROMOTES ITS OWN DEGRADATION IN VIVO. IT
APPEARS THAT ONLY UNMODIFIED E6-AP MOLECULES CAN BIND EFFICIENTLY
TO P53 IN THE PRESENCE OF THE HPV E6 ONCOPROTEIN.
-!- SUBCELLULAR LOCATION: Nuclear (Probable).
-!- ALTERNATIVE PRODUCTS:
Event-Alternative splicing; Named isoforms=3;
Name=II;
IsoId=Q05086-1; Sequence=Displayed;
Name=I;
IsoId=Q05086-2; Sequence=VSP_006705;
Name=III;
IsoId=Q05086-3; Sequence=VSP_006706;
-!- DISEASE: DEFECTS IN UBE3A ARE A CAUSE OF ANGELMAN SYNDROME (AS),
ALSO KNOWN AS 'HAPPY PUPPET SYNDROME'. IT IS CHARACTERIZED BY
FEATURES OF SEVERE MOTOR AND INTELLECTUAL RETARDATION,
MICROCEPHALY, ATAXIA, FREQUENT JERKY LIMB MOVEMENTS AND FLAPPING
OF THE ARMS AND HANDS, HYPTONIA, HYPERACTIVITY, HYPOPLASMATION,
SEIZURES, ABSENCE OF SPEECH, FREQUENT SMILING AND EPISODES OF
PAROXYSMAL LAUGHTER, AND AN UNUSUAL FACIES CHARACTERIZED BY
MACROSTOMIA, A LARGE MANDIBLE AND OPEN-MOUTHED EXPRESSION, A GREAT
PROPENSITY FOR PROTRUDING THE TONGUE ('TONGUE THRUSTING'), AND AN
OCCIPITAL GROOVE.
-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
THIOESTER FORMATION.
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-!- SIMILARITY: SOME, IN THE C-TERMINAL HALF, TO RAT 100 kDa PROTEIN.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
or send an email to license@ebi.ac.uk).

EMBL; X98021; CAA66653.1; JOINED.
EMBL; X98027; CAA66653.1; JOINED.
EMBL; X98022; CAA66653.1; JOINED.
EMBL; X98023; CAA66653.1; JOINED.
EMBL; X98024; CAA66653.1; JOINED.
EMBL; X98025; CAA66653.1; JOINED.
EMBL; X98026; CAA66653.1; JOINED.
EMBL; X98028; CAA66653.1; JOINED.
EMBL; X98029; CAA66653.1; JOINED.
EMBL; X98030; CAA66653.1; JOINED.
EMBL; X98033; CAA66656.1; -.
EMBL; X98031; CAA66654.1; -.
EMBL; X98032; CAA66655.1; -.

EMBL; L07557; AAA35542.1; ALT_INIT.
EMBL; AF016708; AAB69154.1; -.
EMBL; AF016703; AAB69154.1; JOINED.
EMBL; AF016704; AAB69154.1; JOINED.
EMBL; AF016705; AAB69154.1; JOINED.
EMBL; AF016706; AAB69154.1; JOINED.
EMBL; AF016707; AAB69154.1; JOINED.
EMBL; U84404; AAB49301.1; -.
PDB; 1C42; 17-NOV-99.
PDB; 1DSF; 17-NOV-99.
PDB; 1EQX; 28-FEB-01.
DR PDB; 1EQX; 28-FEB-01.
DR Genew; HGNC:12496; UBE3A.
DR MIN; 601623; -.
DR MIN; 105830; -.
GO; GO:0004840; Fubiquitin conjugating enzyme activity; TAS.
GO; GO:0007420; P:brain development; TAS.
GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
DR InterPro; IPR000569; HECT_domain.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS0237; HECT; 1.
KW Nuclear protein; ubl conjugation pathway; Ligase;
KW Alternative splicing; Disease mutation; Polymorphism; 3D-structure.
FT DOMAIN 394 399 ASP/GLU-RICH (ACIDIC).
FT BINDING 776 875 HECT.
FT BINDING 843 843 UBIQUITIN (BY SIMILARITY).
FT BINDING 1 23 MISSING (in isoform I).
FT VARSPLIC 1 10 MEKHCQYK -> MATAKCR (in isoform III).
FT VARSPLIC 1 10 /FTId=VSP_006706.
FT VARIANT 44 44 /FTId=VSP_006706.
FT VARIANT 62 62 C -> Y (PROBABLY POLYMORPHISM).
FT VARIANT 201 201 /FTId=VAR_007852.
FT VARIANT 290 290 /FTId=VAR_008142.
FT VARIANT 372 372 /FTId=VAR_008142.
FT VARIANT 826 826 /FTId=VAR_007853.
FT VARIANT 359 359 G -> V.
FT VARIANT 423 423 S -> P.
FT VARIANT 447 447 I -> II (IN AS).
FT VARIANT 469 469 R -> RNLVNEFSR (IN REF. 2).
FT CONFLICT 423 423 P -> L (IN REF. 2).
FT CONFLICT 647 649 TFR -> LFV (IN REF. 2).
FT CONFLICT 669 669 E -> V (IN REF. 2).
FT CONFLICT 686 686 D -> N (IN REF. 2).
SQ SEQUENCE 875 AA; 100645 MW; 667925F1956670AA CRC64;
Query Match 13.3%; Score 539; DB 1; Length 875;
Best Local Similarity 25.1%; Pred. No. 1.1e-20;
Matches 183; Conservative 117; Mismatches 277; Indels 152; Gaps 22;
QY 105 NLRADTGYQLDLCKLGNPNDVTRGQIVVLSQSDRIGTG-----GQVVDGS 152
D 212 SRIGDSQGNLQKLGPDVSDIDAIR-RVYTRLLSNEKIEAFNALVYLSNPVCD 270
QY 153 RLFPNDLPDNGHEERTASGQIYVNIHTTQWERTPRASEYSSSPGRLSCFVDENTPI 212
D 271 LTYEN-----VYSRDPNYLNF--IIGMENRNLHSPEYLEMALPLFCRAMSKLPL 318
QY 213 SGT-----NGATCGQSDPRLAERR----- 232
D 319 AAOGLRLMSKYNADQIRRMETFFQOLITYKISNEFNSRLVNDDDAIVAASKLKMV 378
QY 233 -----VRSQRHNTMSRTHLHTPPDLP-----GYEORTTQOGQVYFHTQTGVSTWH 280
D 379 YYANVWGGEVDTHNNEEDDEEPIPESELTLQELLGEERNNKKGPRVDFLTGELGVKT-- 436
QY 281 DPRVPRDLNSINCELGFLPPGWEIRNTATGRV-----YFVDHNNRTTQFTDPRIL- 330
D 437 -----LDCKK--PLIPFEFNEPLNEVLEMDKDYTFKVTENKFSFMTCTFFL 484

523 SLVQGOQLNPFYTLKVRROHII DDLALVRLLEMIAMENPADLKKQLYVEPEGQVDEGVS 582
439 REMWYLLSHMLNPNYGLFOYSDDDIYTL-QINPDSAVNPEHLSHFVGRIMGMAVPHG 497
583 KEFFQLVVEIFNPDIQGMFTY--DEATKLFWNPSSF---ETEGQFTLIG-ILGLAIYNN 636
498 HYIDGGFTLPFYKOLIGKSTLDDMELVDPDLHNSLVWILEND-----ITGVLDHTF 549
637 CILDVHPVWVYKLMKGGKTFDLDGSHPLVQSLKDLLEYEGSVEDDMITFOISOT- 695
550 CVENHAYGEIIQHLEHLPNGKSIIPVNEENKVEYVLYVNRFLRGIEAQLAQKGFNEVI 609
696 ----DLFGNPMYDLKENGDKBITNENKKEFVSYSDYILNKSVKQPKAPRRGFHMT 751
610 POHLKLTPEKELELIICGLKIDVNDKVNRLK-HCTPDNIVKMFKAVERFEDEER 667
752 NESPLKLFPEIEHLLICSRNLDFOALETTEYDGGVTRSVIRFELVHSFTDEQ 811
668 RARLLQVTVSSRPVLOG----FKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRIDIPP 724
812 KRLFLLETTGTDRAPVGLGKLVITAKNGPD-----TERLPTSHTCFNVLLLPE 861
725 YESYEKLYEKLITAIETCCFAV 747
862 YSSKEKLVNRLKAITYAKFGM 884

RESULT 12
HUL4_YEAST STANDARD; PRT; 892 AA.
P40985;
01-FEB-1995 (Rel. 31, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Probable ubiquitin--protein ligase HUL4 (EC 6.3.2.-):
HUL4 OR YR036C OR J1608 (Baker's yeast).
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
STRAIN=S288C;
MEDLINE=95397595; PubMed=7668047;
Huang M.-E., Chuat J.-C., Galibert F.;
"Analysis of a 42.5 kb DNA sequence of chromosome X reveals three
tRNA genes and 14 new open reading frames including a gene most
probably belonging to the family of ubiquitin-protein ligases";
Yeast 11:775-781(1995).
[2]
SEQUENCE OF 362-892 FROM N.A.
MEDLINE=95045431; PubMed=7957102;
van Gool A.J., Verhage R., Swagemakers S.M.A., van de Putte P.,
Brouwer J., Troelstra C., Bootsma D., Hoeijmakers J.H.J.;
"RAD26, the functional S. cerevisiae homolog of the Cockayne syndrome
B gene ERCC6";
EMBO J. 13:5361-5369(1994).
[3]
GENE NAME, AND GENE DISRUPTION.
MEDLINE=99077972; PubMed=9858558;
Wang G., Yang J., Huibregtse J.M.;
"Functional domains of the rps ubiquitin-protein ligase";
Mol. Cell. Biol. 19:342-352(1999).
-!- FUNCTION: PROBABLE E3 UBIQUITIN-PROTEIN LIGASE WHICH ACCEPTS
UBIQUITIN FROM AN E2 UBIQUITIN-CONJUGATING ENZYME IN THE FORM OF A
THIOESTER AND THEN DIRECTLY TRANSFERS THE UBIQUITIN TO TARGETED
SUBSTRATES. NON ESSENTIAL.
-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
THIOESTER FORMATION (BY SIMILARITY).
-!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L36344; AAA88738.1; -;
CC EMBL; Z49536; CA889563.1; -;
CC EMBL; X81635; CAAS7291.1; -;
CC PIR; S57055; S57055.
CC SGD; S0003797; YR036C.
CC InterPro; IPR000569; HECT_domain.
CC Pfam; PF00632; HECT; 1.
CC SMART; SM00119; HECTC; 1.
CC PROSITE; PS50237; HECT; 1.
CC Ubl conjugation pathway; Llgase.
CC DOMAIN 792 892 HECT.
CC FT BINDING 860 860 FT
CC FT CONFLICT 362 362 H -> Q (IN REF. 2).
CC FT CONFLICT 410 410 T -> M (IN REF. 2).
CC FT CONFLICT 418 418 T -> I (IN REF. 2).
CC FT CONFLICT 514 519 GKSDV -> QIIRER (IN REF. 2).
CC FT CONFLICT 589 607 ESSRSWFAIDPNFDSKSG -> KKAHVHGLPLTILQILTNQ
CC FT CONFLICT 705 705 R (IN REF. 2).
CC FT CONFLICT 723 724 V -> A (IN REF. 2).
CC FT CONFLICT 723 724 KH -> ND (IN REF. 2).
CC SQ SEQUENCE 892 AA; 103456 MW; 35EF9A8DDA92BD84 CRC64;

Query Match 11.4%; Score 460.5; DB 1; Length 892;
Best Local Similarity 24.5%; Pred. No. 4,7e-25;
Matches 172; Conservative 114; Mismatches 266; Indels 151; Gaps 25;

Qy 96 CVELLSNAINRLKDTGYRLDCKLPGNDNDTVRQIV---VSLQSDRIQTGGVDCS 152
Db 285 CIGVLSNLSTK--TSQQLIQSLRRTPTDNFSYQVEILNLVNFQ-----FS 328
Qy 153 RLFDNDLPQWEERRATASRIQYLNHITRTQWERPTPASEYSPGRLSCFVDENTPI 212
Db 329 RLLSNEL-----SNRTAKNVKPEDEM--RSRLRHHTTGHEFLST-RPIS- 371
Qy 213 SGTNGATCGQSDPRLA--ERRVRSQ---RHNVMNSTRHLHTPPDLPEGYRQTQQGV 267
Db 372 -----AQSDNKGSGFTHPVNNKMKFKFPQYEDNHHSAAKLFTIY----- 413
Qy 268 YFLHTGTGVSTWHDPRVPRDLNINCELGPLPPGWEIRNTATGRVYFVDHNRRTQTD 327
Db 414 YVANT-----RRNGRGALSISQSFYNTLDFID 440
Qy 328 PRLSANLH---LVLNRONLKQQQQVWSLCPDDTECLTVPRYKRDVLQKLI----- 378
Db 441 YKQDFD-HRWGVAQKTKMQLIEWGNSTTK-----KCFSECKYFFLLSLGKISIMEY 493
Qy 379 -LRQELSQQQPQA-----GHCRIVSRBEIPEESYRQVQKMRPKDLWRLMIK 425
Db 494 EIRIMEHEAEQAFILSLDKGKSDVDYFKIKVRDVI SHDSL R-CIKEHQGDLLKSLRIE 552
Qy 426 FRCEGLDYGVAREWLYLLSHEMLNPFYGLFQYSRDDIYT-LQINP---DSAVNPEHL 480
Db 553 FVNEPGIDAGLRKEWFFLLTKSLFNPVNGLFYIKESRSWFAIDPNFDSKSGNSQL 612
Qy 481 SYFHFVGRINGMAVPHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWIL--- 537
Db 613 ELVYLFQVWGLAIFNSTILDLPFKALYKLCSEPLSFEDYSELPFETSRNLKMLNYT 672
Qy 538 ENDITGVLDHTF-----CVEHNAIGEIIQHLEHLPNGKSIIPVNEENKVEYVLYVN 587
Db 673 EDNPFDPVSLTFETTYRNANNWILNDSKSKSEYVTVVELCENGKRVITQSNKHEFVKNWE 732
Qy 588 WRFLRGIEAQLAQKGFNEVIPO-HLAKTFDEKELELIICGLKIDVNDK- ---VNTR 642
Db 733 FYLEKSIETQYNKVFSGFKRVFAECNSIKLFNSELERLVCGDEBQTKFDFKSLRSVTKY 792

```
643 LKCTPDNSIVKWFKAVERFFDEERARLLQFVTGSSRVLQGFALQGAAGPRLFTIHQ 702
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
793 VGFSDSDRAVCWFELISWDYFLOKLLQFVTASDRIPATGISTIFPKIS--LLGSHD 850
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
703 IDACTNNLPKHAHCFNRIDIPPPYSEYKELKLLTAIBETQGF 745
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
851 SD----DLPLAHTCFNEICLWNYSSKKLEKLLWAINESGY 889
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

SULT 13
R3_HUMAN STANDARD; PRT; 1050 AA.
Q15034;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
HECT domain and RCC1-like domain protein 3.
HERC3 OR KIAA0032.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Bone marrow;
MEDLINE=96051387; PubMed=7584026;
Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayasi Y.,
Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
"Prediction of the coding sequences of unidentified human genes. I.
The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by
analysis of randomly sampled cDNA clones from human immature myeloid
cell line KG-1."; (1994).
DNA Res. 1:27-35 (1994).
[2]
CHARACTERIZATION.
MEDLINE=21099818; PubMed=11163799;
Cruz C., Ventura F., Bartrons R., Rosa J.L.;
"HERC3 binding to and regulation by ubiquitin.";
FEBS Lett. 488:74-80 (2001).
-1- SUBCELLULAR LOCATION: Cytoplasmic. Also found in vesicular-like
structures.
-1- PTM: Substrate of ubiquitination and is degraded by the
proteasome.
-1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-1- SIMILARITY: Contains 7 RCC1 repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; D25215; BAA04945.1; -.
Genew; HGNC:4876; HERC3.
MIM; 605200; -.
InterPro; IPR000569; HECT domain.
InterPro; IPR000408; Reg_chr_condens.
Pfam; PF00632; HECT; 1.
Pfam; PF00415; RCC1; 4.
PRINTS; PR00633; RCCNSDMSATION.
SMART; SM00119; HECTC; 1.
PROSITE; PS02037; HECT; 1.
PROSITE; PS00625; RCC1_1; FALSE_NEG.
PROSITE; PS00625; RCC1_2; 4.
PROSITE; PS00625; RCC1_3; 4.
PROSITE; PS00625; RCC1_4; 4.
PROSITE; PS00625; RCC1_5; 4.
PROSITE; PS00625; RCC1_6; 4.
PROSITE; PS00625; RCC1_7; 4.
Ubl conjugation pathway; Ubl conjugation; Repeat.
REPEAT
1 51
52 101
53 154
54 207
REPEAT
102 154
REPEAT
156 207
REPEAT
```

```
FT REPEAT 208 259 RCC1 5.
FT REPEAT 261 311 RCC1 6.
FT REPEAT 313 366 RCC1 7.
FT DOMAIN 951 1050 HECT.
FT BINDING 1018 1018 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 1050 AA; 117188 MW; 5F08A1DE1F40B912 CRC64;

Query Match 11.0%; Score 445.5; DB 1; Length 1050;
Best Local Similarity 31.6%; Pred. No. 6.9e-24;
Matches 115; Conservative 69; Mismatches 151; Indels 29; Gaps 9;

QY 395 IEVSREIFESYQVWMPKPDWKLMIKFRGEEGLDYGVAREWLYLLSHEMLNPPY 454
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 704 LHVRRNNLVGDALRELSHSDIDLKPLKVFDEEAVDAGGVTKPEFLLLKELLNPIY 763
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 455 GLFOYSRDDIYTLQINPDSAVNPSELSYFHEVGRIMGMAVFHGHYIDGGFTLPYKOLLG 514
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 764 GMFTYQDS--NLLWFSDTCP-VEH-NWFHLIGTICGLAINSTVVDLHFLPLAAYKKLN 819
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 515 KSITLDDMELVDPDLHNSLWLENDITGV-LDHTFCVE----HNAYGEIIQHLEKNGK 569
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 820 VKPGLDLKELSPTEGRSLQELL--DYPGEDVETPCLNFTICRESYGVIEQKKLIPEGD 877
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 570 SIPVNEENKKEVRLVYVNWRLRGIEAQLQGFNEVTPQHLKTFDEKLELLICGL 629
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 878 NTVCKONRQEFVDAYVNYVFOISVHEWYTAFFSGFLKVGCGKVLQFQPSSELRMMVGN 937
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 630 GKIDVND-----WKVNTRLKCHTPDSDNIVKWFKAVERFFDEERARLLQFVTGSSRVL 683
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 938 SNYNWEELEETAIYKGYDSATHPT-----VKLWETTFEPLEKKKKFLFLTGSDRIPI 992
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 684 QGFALQGAAGPRLFTIHOIDACTNNLPKHAHCFNRIDIPPPYSEYKELKLLTAIBETC 743
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 993 YGMASLQ-----IVIOSTASGEYLPVVAHTCYNLLDLPKYSKEILSARLTQALDNYE 1045
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 744 GFAP 747
      :  :  :
Db 1046 GFSL 1049

RESULT 14
TRIB_HUMAN STANDARD; PRT; 1992 AA.
AC Q14669; Q15644;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Thyroid receptor interacting protein 12 (TRIP12).
GN TRIP12 OR KIAA0045.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawarabayasi Y., Ishikawa K.-I., Tabata S.;
RA "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RN DNA Res. 1:223-229 (1994).
[2]
RP SEQUENCE OF 1801-1992 FROM N.A.
RX MEDLINE=95295737; PubMed=7776974;
RA Lee J.W., Choi H.-S., Gyuris J., Brent R., Moore D.D.;
RA "Two classes of proteins dependent on either the presence or absence
RT of thyroid hormone for interaction with the thyroid hormone
RT receptor.";
RN Mol. Endocrinol. 9:243-254 (1995).
CC -1- FUNCTION: COMPONENT OF PA700, AN ATP-DEPENDENT MULTISUBUNIT
CC PROTEIN THAT ACTIVATES THE PROTEOLYTIC ACTIVITIES OF THE
```


180	TRT-TOWERPTPASEYSPGRPLSCFVDENTISGTNG--ATCQSSDDPRLAERRVRSQ	236
1391	TKTHITWYKPVFVDEESNK-----DC-----VGKRGRAQTAFTKTSPRNAK-----	1433
237	RHNYNSRTHLHTPPDLPEGVEQRTTOQGOVYLHTQTGVSTWHPD	288
1434	---HDELWHGVCFSVSNPLE-----VYLPTPENITFEDPSLDVILLRLVLHAI	1481
289	S-----NINCEBGLPLPGWEIRNATGRVYVDHNNRTQTTPRLSANLHLVNR	340
1482	SRWVYLYDNAMCKEILP-----TSEFINSKLTAKAN-----	1513
341	QNOLKD-----OQOOQVNSLCP-----DDECL-TVPRYKRD-LVQKGLKILSQEL	383
1514	-ROLOQPLVIMTGNITPTWLTGELTCFPFFPTQMLFYVTAPODRAMQRLDWTNPEI	1572
384	SOQQ-----POAGCHRIEVSREIEESYOVMMQRPKDLW-KRLMIKRGREGLDYQ	435
1573	NQSDSDSRVAPRLDKRTVNRBELUKQA-ESVM-----ODLGSSRAMLEIQYENEVGTG	1627
436	-GVAREWLYLLSHEMLNPYGLFCYSDDITTIQINPDSA-----	474
1628	LGTLEFYALVNSQELQORADLGLW---RGEVTLIS-NPKXSQEGTKYVIONLOQLFALPFGR	1683
475	-VNPEHLS-----YHFVGRINGMAVFHGHYDGGFTLPFYKOLLAG--SITLDDMELVDP	527
1684	TAKPAHIAKVKVKFFLGLKMAKAIMDFRVLVDPLGLPFYKVMLORETSLSHOLFIDP	1743

